Contents

1

Compressive Optical Imaging: Architectures and Algorithms

Roummel F. Marcia, Rebecca M. Willett, and Zachary T. Harmany [1]

1.1
Introduction

Many traditional optical sensors are designed to collect directly interpretable and intuitive measurements. For instance, a standard digital camera directly measures the intensity of a scene at different spatial locations to form a pixel array. Recent advances in the fields of image reconstruction, inverse problems, and compressive sensing (CS) [1, 2] indicate, however, that substantial performance gains may be possible in many contexts via less direct measurements combined with computational methods. In particular, CS allows for the extraction of high-resolution images from relatively small focal plane arrays (FPAs). CS is described in detail in Chapter 23. The basic idea of CS theory is that when the image of interest is very sparse or highly compressible in some basis (i.e., most basis coefficients are small or zero-valued), relatively few well-chosen observations suffice to reconstruct the most significant non-zero components. In particular, judicious selection of the type of image transformation introduced by measurement systems may dramatically improve our ability to extract high-quality images from a limited number of measurements. By designing optical sensors to collect measurements of a scene according to CS theory, we can use sophisticated computational methods to infer critical scene structure and content.

These ideas are particularly relevant to imaging applications in which it is useful or even crucial to keep the FPA of an optical system relatively small. For example, in low light settings where sensitive detectors are costly, smaller FPAs translate directly to less expensive systems. Smaller FPAs also make systems lighter weight and thus more portable. In addition, imaging systems with fewer photodetectors consume less power

and therefore require fewer battery charges. Finally, smaller cameras can fit into tighter spaces for unobtrusive surveillance. An important goal in the design of many imaging systems, then, is to extract as much information as possible from a small number of detector array measurements.

While recent progress in the exploitation of CS theory is highly encouraging, there are several key issues in the context of optical systems that must be addressed:

- In most real-world sensing systems, we face **physical constraints** on the nature of the scene (i.e., photon intensity is always nonnegative) and the nature of the measurements which can be collected in hardware (i.e., photons can be accumulated, redirected, or filtered, but not "subtracted" in conventional settings).

- A typical CS assumption is that each measurement collected by a sensor is the projection of the image of interest onto a different random vector. It is not clear how **practical optical systems** are best built within this context. In many settings, collecting these projections would result in either a physically very large and cumbersome system or significant noise when sequences of measurements are collected over a limited amount of time.

- It typically is not possible to wait hours or even minutes for an iterative reconstruction routine to produce a single image; rather, algorithms must be able to operate effectively under **stringent time constraints** and produce image estimates which satisfy the above physical constraints.

In this chapter, we explore (a) the potential of several different physically realizable optical systems based on CS principles and (b) associated fast numerical reconstruction algorithms for overcoming the challenges described above. Specifically, for a fixed size FPA, we describe how compressive measurements combined with sophisticated optimization algorithms can significantly increase image quality and/or resolution.

Organization of chapter. The chapter is organized as follows. We describe the compressive sensing problem and formulate it mathematically in Section 1.2. Section 1.3 describes architectures for compressive image acquisition. In particular, we focus on coded aperture imaging and how CS theory can be used for constructing coded aperture masks that can easily be implemented for improving image reconstruction resolution. Section 1.4 discusses algorithms for recovering compressively sensed images while incorporating physical constraints inherent to optical imaging. In addition, we describe a new recovery algorithm which can exploit structure in sparsity patterns. We present experimental results and provide analysis in Section 1.5, discuss the impact of noise and quantization errors in Section 1.6, and offer concluding remarks in Section 1.7.

1.2

Compressive Sensing

To understand the intuition behind the CS framework, consider a $\sqrt{n} \times \sqrt{n}$ image $\boldsymbol{G}$, which can be written as a length-n column vector $\boldsymbol{g}$ and represented in terms of a basis expansion with n coefficients:

$$\boldsymbol{g} = \sum_{i=1}^{n} \theta_i \boldsymbol{w}_i,$$

where $\boldsymbol{w}_i$ is the ith basis vector and θ_i is the corresponding coefficient. In many settings, the basis $\boldsymbol{W} \triangleq [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n]$ can be chosen so that only $k \ll n$ coefficients have significant magnitude, i.e., many of the θ_i's are zero or very small for large classes of images; we then say that $\boldsymbol{\theta} \triangleq [\theta_1, \ldots, \theta_n]^t$ is *sparse* or *compressible*. In such cases, it is clear that if we knew *which k θ_i's* were significant, we would ideally just measure these k coefficients directly, resulting in fewer measurements to obtain an accurate representation of $\boldsymbol{g}$. Of course, in general we do not know *a priori* which coefficients are significant. The key insight of CS is that, with slightly more than k well-chosen measurements, we can determine which θ_i's are significant and accurately estimate their values. Furthermore, fast algorithms which exploit the *sparsity* of $\boldsymbol{\theta}$ make this recovery computationally feasible. Sparsity has long been recognized as a highly useful metric in a variety of inverse problems, but much of the underlying theoretical support was lacking. However, more recent theoretical studies have provided strong justification for the use of sparsity constraints and quantified the accuracy of sparse solutions to these underdetermined systems [1, 3].

The problem of estimating the image $\boldsymbol{G}^{\text{true}} \in \mathbb{R}_+^{\sqrt{n} \times \sqrt{n}}$, where n is the total number of image pixels or voxels, can be formulated mathematically as an inverse problem, where the data collected by an imaging or measurement system are represented as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{g}^{\text{true}} + \boldsymbol{n}, \tag{1.1}$$

where $\boldsymbol{g}^{\text{true}} \in \mathbb{R}_+^n$ is the image $\boldsymbol{G}^{\text{true}}$ stored as a length-n column vector of pixel intensities, $\boldsymbol{A} \in \mathbb{R}_+^{m \times n}$ linearly projects the scene onto a m-dimensional set of observations, $\boldsymbol{n} \in \mathbb{R}^m$ is noise associated with the physics of the sensor, and $\boldsymbol{y} \in \mathbb{R}_+^m$ is the observed data. (Typically $\boldsymbol{n}$ is assumed bounded or bounded with high probability to ensure that $\boldsymbol{y}$, which is proportional to photon intensities, is nonnegative.) CS addresses the problem of solving for $\boldsymbol{g}^{\text{true}}$ when the number of unknowns, n, is much larger than the number of observations, m. In general, this is an ill-posed problem as there are an infinite number of candidate solutions for $\boldsymbol{g}^{\text{true}}$; nevertheless, CS theory provides a set of conditions that, if satisfied, assure an accurate estimation of $\boldsymbol{g}^{\text{true}}$. We first presuppose that $\boldsymbol{g}^{\text{true}}$ is sparse or compressible in a basis $\boldsymbol{W}$. Then given $\boldsymbol{W}$, we require that $\boldsymbol{A}$ in conjunction with $\boldsymbol{W}$ satisfies a technical condition called the *Restricted Isometry Property* (RIP) [1]. More specifically, we say that $\boldsymbol{A}\boldsymbol{W}$ satisfies the RIP of order s if there exists a constant $\delta_s \in (0, 1)$ for which

$$(1 - \delta_s)\|\boldsymbol{\theta}\|_2^2 \leq \|\boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2 \leq (1 + \delta_s)\|\boldsymbol{\theta}\|_2^2. \tag{1.2}$$

holds for all s-sparse $\boldsymbol{\theta} \in \mathbb{R}^n$. In other words, the energy contained in the projected image, $\boldsymbol{AW\theta}$, is close to the energy contained in the original image, $\boldsymbol{\theta}$. While the RIP cannot be verified for an arbitrary given observation matrix and basis, it has been shown that observation matrices $\boldsymbol{A}$ drawn independently from many probability distributions satisfy the RIP of order s with high probability for any orthogonal basis $\boldsymbol{W}$ when $m \geq Cs \log(n/s)$ for some constant C [1]. Although generating CS matrices using this procedure is simple in software, building physical systems to measure $\boldsymbol{AW\theta}$ can be notoriously difficult.

While the system of equations in (1.1) can be grossly underdetermined, CS theory suggests that selecting the *sparsest* solution to this system of equations will yield a highly accurate solution, subject to $\boldsymbol{\theta}^{\mathrm{true}} = \boldsymbol{W}^t \boldsymbol{g}^{\mathrm{true}}$ being sufficiently sparse and $\boldsymbol{A}$ satisfying the RIP of order $2k$, where k is the sparsity of $\boldsymbol{g}^{\mathrm{true}}$ [1]. This problem can be posed as the optimization problem

$$\boldsymbol{\theta}^{\mathrm{est}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_0 \quad \text{subject to} \quad \|\boldsymbol{y} - \boldsymbol{AW\theta}\|_2 \leq \epsilon, \tag{1.3a}$$

$$\boldsymbol{g}^{\mathrm{est}} = \boldsymbol{W}\boldsymbol{\theta}^{\mathrm{est}} \tag{1.3b}$$

where the ℓ_0 quasinorm $\| \cdot \|_0$ denotes the number of nonzero elements in the argument and $\epsilon \geq 0$ is a bound on the noise. In general ℓ_0 minimization is NP-hard since it is combinatorial in nature and its computational complexity grows exponentially with the size of the image [4]. A common alternative is to *relax* this norm to the ℓ_1 norm [2, 5, 6], and instead solve

$$\boldsymbol{\theta}^{\mathrm{est}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - \boldsymbol{AW\theta}\|_2 \leq \epsilon, \tag{1.4a}$$

$$\boldsymbol{g}^{\mathrm{est}} = \boldsymbol{W}\boldsymbol{\theta}^{\mathrm{est}}. \tag{1.4b}$$

One of the key insights of CS is that in the noiseless case, i.e., when the bound ϵ is set to zero, the solution to this relaxed problem is *equivalent* to the solution of the intractable problem in (1.3) if the matrix product $\boldsymbol{AW}$ satisfies the RIP (1.2). We discuss equivalent formulations of (1.4) and various ways of solving them in Sec. 1.4.

1.3
Architectures for Compressive Image Acquisition

Developing practical optical systems to exploit CS theory is a significant challenge being explored by investigators in the signal processing, optics, astronomy, and coding theory communities. In addition to implicitly placing hard constraints on the nature of the measurements which can be collected, such as nonnegativity of both the projection vectors and the measurements, practical CS imaging systems must also be robust and reasonably sized. In [7], Neifeld and Ke describe three general optical architectures for compressive imaging: (1) sequential, where measurements are taken one at a time, (2) parallel, where multiple measurements taken simultaneously using a fixed mask, and (3) photon-sharing, where beam-splitters and micromirror arrays are used to collect
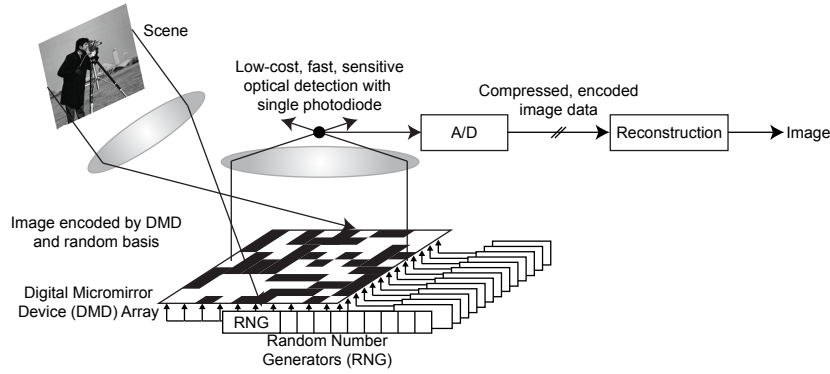
Fig. 1.1 A block diagram of the Rice single-pixel camera. Light from the scene to be imaged is reflected off a digital micromirror device (DMD) array whose mirror pattern orientations are generated (pseudo)randomly. Each pattern produces a single measurement. From a collection of measurements, an estimate of the scene is constructed using compressive sensing (CS) algorithms.

measurements. Here, we describe some optical hardware with these architectures that have been recently considered in literature.

Rice single-pixel camera. The Rice single-pixel camera [8] is a camera architecture that uses only a single detector element to image a scene. As shown in Fig. 1.1, a digital micromirror array is used to represent a pseudo-random binary array, and the scene of interest is then projected onto that array before the aggregate intensity of the projection is measured with a single detector. Since the individual orientations of the mirrors in the micromirror array can be altered very rapidly, a series of different pseudorandom projections can be measured successively in relatively little time. The original image is then reconstructed from the resulting observations using CS reconstruction techniques such as those described in Section 1.4. One of the chief benefits of this architecture is that any binary projection matrix can readily be implemented in this system, so that existing CS theory can be directly applied to the measurements. While this "single-pixel" imaging system demonstrates a successful implementation of CS principles, the setup limits the temporal resolution of a dynamically changing scene. Although we can rapidly collect many scenes sequentially at lower exposure, this increases the amount of noise per measurement, thus diminishing its potential for video imaging applications.

Spectral imagers. In spectral imaging settings, a vector containing spectral information about the materials being imaged is assigned to each spatial pixel location, forming a three-dimensional data cube (i.e., two spatial dimensions and one spectral dimension). However, tradeoffs between spectral and spatial resolution limit the performance of modern spectral imagers, especially in photon-limited settings where the small number of photons must be apportioned between the voxels in the data cube, resulting in low signal-to-noise ratio per voxel. In [9], investigators propose innovative, real-time spectral imagers, where each pixel measurement is the coded projection of the spectrum in the corresponding spatial location in the data cube. This was implemented

using two dispersive elements separated by binary-coded masks. In numerical experiments, a $256 \times 256 \times 15$ spectral image was accurately recovered from a 256×256 observation using an appropriate reconstruction approach. In related work, objects are illuminated by light sources with tunable spectra using spatial light modulators to facilitate compressive spectral image acquisition [10].

Task-specific information compressive imagers. Compressive optical architectures have been considered for target detection systems which use knowledge of the target's features to reduce the number of measurements required from a more conventional snapshot camera. In particular, investigators in [11] note that effective target detection can be performed using a relatively small subset of coefficients of principal components, independent components, generalized matched-filters, or generalized Fisher discriminant representations; as a result, directly measuring these coefficients can be more efficient than collecting direct pixel-wise observations of the entire scene and then computing these coefficients in software. The authors further describe ways to optimize the amount of photon energy allotted to each representation element (e.g., each principal component measured) in order to maximize target detection performance. While the proposed architectures are finely tuned to specific target detection tasks and cannot be used for some general purpose imaging tasks, they nevertheless carefully consider the role of physical constraints introduced by optical elements and photon noise.

Coded aperture imagers. Investigators in [12] and [13] propose practical implementations of CS ideas using coded apertures, demonstrating that if the coded apertures are designed using a pseudorandom construction, then the resulting observation model satisfies the RIP. This approach is described in detail in the next section. This parallel architecture is highly suitable for practical and implementable compressive imaging since it provides a snapshot image (i.e., all m measurements are collected simultaneously) and does not require complex, and potentially large, imaging apparatuses.

1.3.1
Coded Apertures

Coded apertures were first developed to increase the amount of light hitting a detector in an optical system without sacrificing resolution (by, say, increasing the diameter of an opening in a pinhole camera). The basic idea is to use a mask, i.e., an opaque rectangular plate with a specified pattern of openings, that allows significantly brighter observations with higher signal-to-noise ratio than those from conventional pinhole cameras [14]. These masks encode the image before detection, and the original image is recovered from the distorted observation in post-processing using an appropriate reconstruction algorithm. The mask pattern introduces a more complicated point spread function than that associated with a pinhole aperture, and this pattern is exploited to reconstruct high-quality image estimates (see Fig. 1.2). These multiplexing techniques are particularly popular in astronomical [15] and medical [16] applications because of their efficacy at wavelengths where lenses cannot be used, but recent work has also

demonstrated their utility for collecting both high resolution images and object depth information simultaneously [17].

Seminal work in coded aperture imaging includes the development of masks based on Hadamard transform optics [18] and pseudorandom phase masks [19]. Modified Uniformly Redundant Arrays (MURAs) [20] are generally accepted as optimal mask patterns for coded aperture imaging. These mask patterns (which we denote by $\boldsymbol{H}^{\mathrm{MURA}}$) are binary, square patterns, whose *grid size matches the spatial resolution of the FPA* and whose sidelength is a prime integer number of grid cells. Each mask pattern is specifically designed to have a complementary pattern $\boldsymbol{H}^{\mathrm{recon}}$ such that $\boldsymbol{H}^{\mathrm{MURA}} * \boldsymbol{H}^{\mathrm{recon}}$ is a single peak with flat side-lobes (i.e., a Kronecker δ function).

In practice, the resolution of a detector array dictates the properties of the mask pattern and hence resolution at which $\boldsymbol{G}^{\mathrm{true}}$ can be reconstructed. We model this effect as $\boldsymbol{G}^{\mathrm{true}}$ being downsampled to the resolution of the detector array and then convolved with the mask pattern $\boldsymbol{H}^{\mathrm{MURA}}$, which has the same resolution as the FPA and the downsampled $\boldsymbol{G}^{\mathrm{true}}$, i.e.,

$$\boldsymbol{Y} = (\boldsymbol{\mathcal{D}}(\boldsymbol{G}^{\mathrm{true}})) * \boldsymbol{H}^{\mathrm{MURA}} + \boldsymbol{N}, \tag{1.5}$$

where $*$ denotes convolution, $\boldsymbol{N}$ corresponds to noise associated with the physics of the sensor, and $\boldsymbol{\mathcal{D}}(\boldsymbol{G}^{\mathrm{true}})$ is the downsampling of the scene, which consists of partitioning $\boldsymbol{G}^{\mathrm{true}}$ into m uniformly sized blocks and measuring the total intensity in each block. (Recall m is the size of the detector array and hence the size of $\boldsymbol{Y}$.) This is sometimes referred to as integration downsampling.

Because of the construction $\boldsymbol{H}^{\mathrm{MURA}}$ and $\boldsymbol{H}^{\mathrm{recon}}$, $\boldsymbol{\mathcal{D}}(\boldsymbol{G}^{\mathrm{true}})$ can be reconstructed using

$$\boldsymbol{G}^{\mathrm{est}} = \boldsymbol{Y} * \boldsymbol{H}^{\mathrm{recon}}.$$

However, the resulting resolution is often lower than what is necessary to capture some of the desired details in the image. Clearly, the estimates from MURA reconstruction are limited by the spatial resolution of the photo-detector. Thus, high resolution reconstructions cannot generally be obtained from low-resolution MURA-coded observations. It can be shown that this mask design and reconstruction result in minimal reconstruction errors *at the FPA resolution* and *subject to the constraint that linear, convolution-based reconstruction methods would be used*.

1.3.2
Compressive Coded Apertures

A recent study by the authors [12] addresses the accurate reconstruction of a high resolution static image which has a sparse representation in some basis from a single low resolution observation using *compressive* coded aperture imaging. In this study, we designed a coded aperture imaging mask such that the corresponding observation matrix $\boldsymbol{A}^{\mathrm{CCA}}$ satisfies the RIP. Here, the measurement matrix $\boldsymbol{A}^{\mathrm{CCA}}$ associated with compressive coded apertures (CCA) can be modeled as

$$\boldsymbol{A}^{\mathrm{CCA}}\boldsymbol{g}^{\mathrm{true}} = \mathrm{vect}(\boldsymbol{\mathcal{D}}(\boldsymbol{G}^{\mathrm{true}} * \boldsymbol{H}^{\mathrm{CCA}})), \tag{1.6}$$
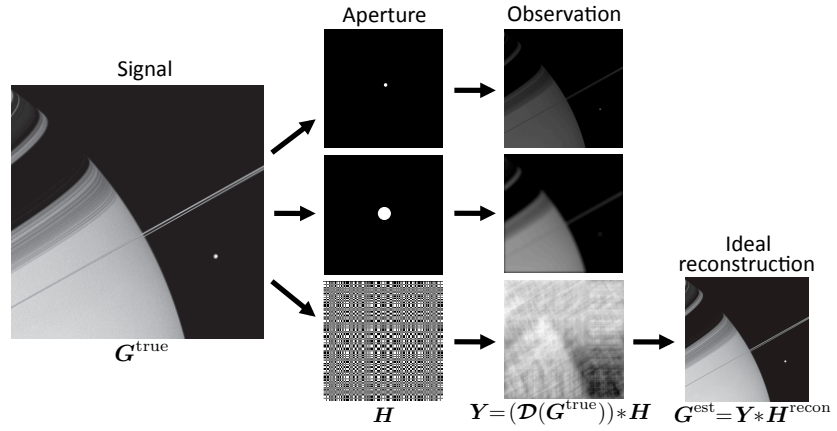
Fig. 1.2 An illustration of aperture imaging. An image G^{true} obtained through a pinhole camera (top) results in dark observations due to the low amount of light allowed through the opening. The aperture can be enlarged (middle) to obtain brighter observations, but this yields blurry observations due to edge effects. Coded apertures, such as MURA patterns H^{MURA} (bottom), allow substantially more light. The observation Y is the convolution between a downsampling of the true image (represented by $\mathcal{D}(G^{\text{true}})$) and H^{MURA}. Although the observation has little resemblance to the original image, a highly accurate estimate G^{est} can be obtained in a post-processing step involving convolution with a decoding pattern H^{recon}.

where H^{CCA} is the coding mask, and $\text{vect}(\cdot)$ is an operator that converts an image into a column vector. Here the coding mask, H^{CCA}, is the size and resolution at which G^{true} will be reconstructed, rather than the size and resolution of the FPA as we had with H^{MURA}. Thus in (1.6), we model the measurements as the scene being convolved with the coded mask and *then* downsampled. In contrast, MURA masks are designed for optimal performance on a downsampled version of the scene, as in (1.5).

The coded aperture masks are designed so that A^{CCA} satisfies the RIP of order $2k$ as described in (1.2) with high probability when $m \geq Ck^3 \log(n/k)$ for some constant C [12, 21]. Note that this is somewhat weaker than what can be achieved with a randomly generated sensing matrix which satisfies (1.2) with high probability when $m \geq Ck \log(n/k)$ for some constant C. The extra factor of k^2 comes from the fact that the m projections sensed using a coded aperture framework are dependent on one another. Stronger results based on a more complex coded aperture architecture were shown recently in [22]. Specifically, if $m \geq C(k \log(n) + \log^3(k))$ for some constant C, then g^{true} can be accurately recovered with high probability. In general, however, to achieve a given level of accuracy, one would need more coded aperture observations than observations from, say, the Rice single-pixel camera. This drawback, however, is mitigated by the fact that coded apertures yield compact and easily implementable snapshot optical designs and relatively little measurement noise for a fixed observation time window.

The convolution of H^{CCA} with an image G^{true} as in (1.6) can be represented as the application of the Fourier transform to G^{true} and H^{CCA}, followed by element-
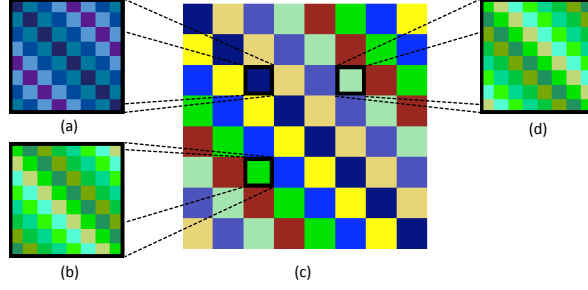
Fig. 1.3 The matrix $\boldsymbol{\mathcal{F}}^{-1}\boldsymbol{C_H}\boldsymbol{\mathcal{F}}$ in (c) is block-circulant with circulant blocks. Enforcing symmetry, i.e., requiring the diagonal blocks (a) to be symmetric and opposing non-diagonal blocks (e.g., those denoted by (b) and (d)) to be transposes of each other, produces a transfer function $\hat{\boldsymbol{H}}^{\mathrm{CCA}}$ that is symmetric about its center, which translates to a real-valued point spread function, and, consequently, a physically realizable coded aperture pattern $\boldsymbol{H}^{\mathrm{CCA}}$.

wise matrix multiplication and application of the inverse Fourier transform. In matrix notation, this series of linear operations can be expressed as

$$\mathrm{vect}(\boldsymbol{H}^{\mathrm{CCA}} * \boldsymbol{G}^{\mathrm{true}}) = \boldsymbol{\mathcal{F}}^{-1}\boldsymbol{C_H}\boldsymbol{\mathcal{F}}\,\boldsymbol{g}^{\mathrm{true}},$$

where $\boldsymbol{\mathcal{F}}$ is the two-dimensional Fourier transform matrix and $\boldsymbol{C_H}$ is the diagonal matrix whose diagonal elements correspond to the transfer function $\hat{\boldsymbol{H}}^{\mathrm{CCA}} \triangleq \boldsymbol{\mathcal{F}}(\boldsymbol{H}^{\mathrm{CCA}})$. The matrix product $\boldsymbol{\mathcal{F}}^{-1}\boldsymbol{C_H}\boldsymbol{\mathcal{F}}$ is block-circulant and each block is in turn circulant (see Fig. 1.3). Block-circulant matrices, whose entries are drawn from an appropriate probability distribution, are known to be CS matrices. Based upon recent theoretical work on Toeplitz- and circulant-structured matrices for CS, the proposed masks are fast and memory-efficient to generate [12, 21, 22]. In addition, the diagonalizability of block-circulant matrices with circulant blocks by the discrete Fourier transform leads to fast matrix-vector products that are necessary for efficient reconstruction algorithms. The incorporation of the integration downsampling operator $\boldsymbol{\mathcal{D}}$ does not prevent the RIP from being satisfied; a key element of the proof that the RIP is satisfied is a bound on the number of rows of $\boldsymbol{A}^{\mathrm{CCA}}$ which are statistically independent. Since the downsampling operator effectively sums rows of a block circulant matrix, downsampling causes the bound on the number of dependent matrix rows to be multiplied by the downsampling factor. Enforcing symmetry on $\boldsymbol{\mathcal{F}}^{-1}\boldsymbol{C_H}\boldsymbol{\mathcal{F}}$ (see Fig. 1.3) is equivalent to assuring that the transfer function matrix $\hat{\boldsymbol{H}}^{\mathrm{CCA}}$ is symmetric about its center, so that the resulting coding mask pattern $\boldsymbol{H}^{\mathrm{CCA}} \triangleq \boldsymbol{\mathcal{F}}^{-1}(\hat{\boldsymbol{H}}^{\mathrm{CCA}})$ will be necessarily real [12]. Contrasting mask patterns for MURA coded aperture imaging vs. compressive coded aperture imaging are displayed in Fig. 1.4.
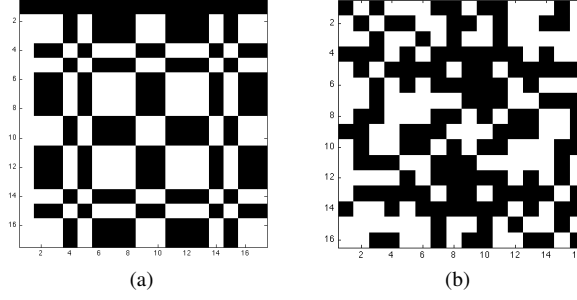
1.4

(a)   (b)

Fig. 1.4 Coded aperture patterns. Here, the white blocks represent the openings in the mask pattern. (a) The 17×17 Modified Uniform Redundant Array (MURA) mask pattern $\boldsymbol{H}^{\mathrm{MURA}}$. The length of the side of the MURA pattern must be a prime number. (b) An example of a 16×16 compressive coded aperture (CCA) mask pattern $\boldsymbol{H}^{\mathrm{CCA}}$. The pattern is not quite random – note the symmetry of the CCA pattern about the $(9, 9)$ element.

Algorithms for Restoring Compressively Sensed Images

The ℓ_1-minimization problem (1.4a) is frequently posed in Lagrangian form:

$$\boldsymbol{\theta}^{\mathrm{est}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2 + \tau\|\boldsymbol{\theta}\|_1, \qquad (1.7)$$

for some regularization parameter $\tau > 0$. This formulation is known as the basis pursuit denoising problem (BPDP) [2], and it has been shown that this ℓ_2-ℓ_1 minimization problem will yield a highly accurate solution of the true underlying image $\boldsymbol{g}^{\mathrm{true}}$ for an appropriate choice of the regularization parameter [3]. In this section, we describe some current methods for solving the relaxed problem (1.7) and its equivalent formulations.

1.4.1
Current Algorithms for Solving the CS Minimization Problem

In order to apply gradient-based methods to solve (1.7), its objective function must be made differentiable since the ℓ_1 norm is not differentiable at 0. In [23], the BPDP is transformed to

$$\operatorname*{minimize}_{\boldsymbol{u}, \boldsymbol{\theta}} \; \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\theta}\|_2^2 + \tau \sum_{i=1}^{n} u_i \quad \text{subject to} \quad -u_i \leq \theta_i \leq u_i, \quad i = 1, \cdots, n,$$

and an interior-point method is applied to obtain the solution. In this approach, iterates are defined in the strict interior of the feasible set, and a logarithmic barrier function is imposed for the bound constraints to prevent the iterates from nearing the boundary of the feasible set prematurely (thus avoiding potentially very small steps along search directions).

An alternative (and often faster) approach to solving (1.7) is the Gradient Projection for Sparse Reconstruction (GPSR) algorithm [24]. In this method, the variable $\boldsymbol{\theta}$ is split into its positive and negative components to make the objective function in (1.7) differentiable. However, this introduces bound constraints on the new variables. GPSR solves this bound constraint optimization problem using a gradient descent method (called gradient projection) that ensures that the iterates remain feasible by projecting (via simple thresholding) onto the feasible region

The Sparse Reconstruction by Separable Approximation (SpaRSA) algorithm [25] reduces (1.7) to a series of alternating steps: (a) approximating the objective function with a regularized quadratic objective, and (b) performing regularized least squares image denoising. This framework allows users to take advantage of the plethora of fast and effective image denoising methods available. More specifically, the SpaRSA approach defines estimates $\{\boldsymbol{\theta}^j\}$ of $\boldsymbol{\theta}^{\text{true}}$ by solving a sequence of quadratic subproblems that approximate (1.7). If $\phi(\boldsymbol{\theta}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2$ in (1.7), then these subproblems are unconstrained quadratic minimization problems of the form

$$\boldsymbol{\theta}^{j+1} = \arg\min_{\boldsymbol{\theta}} \ \phi(\boldsymbol{\theta}^j) + (\boldsymbol{\theta} - \boldsymbol{\theta}^j)^t \nabla\phi(\boldsymbol{\theta}^j) + \frac{\alpha_j}{2}\|\boldsymbol{\theta} - \boldsymbol{\theta}^j\|_2^2 + \tau\|\boldsymbol{\theta}\|_1, \qquad (1.8)$$

where the first three terms in the objective function correspond to the Taylor expansion of $\phi(\boldsymbol{\theta})$ at the current iterate $\boldsymbol{\theta}^j$ and where the second derivative of $\phi(\boldsymbol{\theta})$ is approximated by a multiple of the identity matrix, namely, $\nabla^2\phi(\boldsymbol{\theta}^j) \approx \alpha_j I$, where α_j is defined as in [24, 25]. This approach is particularly fast and effective for RIP-satisfying CS matrices since the near-isometry condition implies that $\boldsymbol{A}^T\boldsymbol{A} \approx \alpha I$. The subproblem (1.8) can be viewed as a denoising subproblem, which, by construction, may have an analytic solution that can easily be calculated. Therefore, the estimates $\{\boldsymbol{\theta}^j\}$ can be computed quickly. We will see this approach of solving an inverse problem by means of a sequence of denoising problems in the subsequent section where we detail how we can adapt these algorithms to the estimation of a light intensity, an inherently nonnegative quantity.

Other methods solve different formulations of the CS minimization problem. A formulation of the CS minimization problem is least absolute shrinkage and selection operator (LASSO) formulation [5] given by

$$\boldsymbol{\theta}^{\text{est}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2 \quad \text{subject to } \|\boldsymbol{\theta}\|_1 \leq \tau.$$

In some instances, a penalty different from the ℓ_1 term is used. The Bregman iterative regularization method [26] uses a *total variation* functional $\tau \int |\nabla\boldsymbol{g}|$ for computing penalties. Finally, we mention (but do not detail for lack of space) two families of algorithms for finding sparse representations: iterative shrinkage/thresholding (IST) algorithms [25] and orthogonal matching pursuits (OMP) [27]. IST methods iteratively map the objective function to a simpler optimization problem which can be solved by shrinking or thresholding small values in the current estimate of $\boldsymbol{\theta}$. OMP methods start with $\boldsymbol{\theta}^{\text{est}} = 0$ and greedily choose elements of $\boldsymbol{\theta}^{\text{est}}$ to have non-zero magnitude by iteratively processing residual errors between $\boldsymbol{y}$ and $\boldsymbol{A}\boldsymbol{\theta}^{\text{est}}$.

1.4.2

Algorithms for Nonnegativity Constrained ℓ_2-ℓ_1 CS Minimization

In optical imaging, we often estimate light intensity, which *a priori* is nonnegative. Thus it is necessary that the reconstruction $\boldsymbol{g}^{\text{est}} = \boldsymbol{W}\boldsymbol{\theta}^{\text{est}}$ is nonnegative, which involves adding constraints to the CS optimization problem (1.7), i.e.,

$$\boldsymbol{\theta}^{\text{est}} = \arg\min_{\boldsymbol{\theta}} \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2 + \tau\|\boldsymbol{\theta}\|_1, \quad \text{subject to} \quad \boldsymbol{W}\boldsymbol{\theta} \geq 0. \quad (1.9)$$

The addition of the nonnegativity constraint in (1.9) makes the problem more challenging than the conventional CS minimization problem, and has only been addressed in CS literature recently in the context of photon-limited compressive sensing [28, 29]. Here, we discuss an approach similar to those in [25, 29] to address the nonnegativity constraints in the ℓ_2-ℓ_1 CS minimization problem.

Let $\phi(\boldsymbol{\theta}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{W}\boldsymbol{\theta}\|_2^2$ be the quadratic term in (1.9). As in the SpaRSA algorithm, the minimization problem (1.9) can be solved using a sequence of quadratic approximation subproblems that are easier to solve. The resulting minimization subproblem is given by

$$\boldsymbol{\theta}^{j+1} = \arg\min_{\boldsymbol{\theta}} (\boldsymbol{\theta} - \boldsymbol{\theta}^j)^t \nabla\phi(\boldsymbol{\theta}^j) + \frac{\alpha_j}{2}\|\boldsymbol{\theta} - \boldsymbol{\theta}^j\|_2^2 + \tau\|\boldsymbol{\theta}\|_1, \text{ subject to } \boldsymbol{W}\boldsymbol{\theta} \geq 0,$$

which is similar to the SpaRSA subproblem (1.8) but with the additional nonnegativity constraint on $\boldsymbol{W}\boldsymbol{\theta}$. This nonnegative denoising subproblem can be written equivalently and more compactly as

$$\boldsymbol{\theta}^{j+1} = \arg\min_{\boldsymbol{\theta}} \quad \frac{1}{2}\|\boldsymbol{s}^j - \boldsymbol{\theta}\|_2^2 + \frac{\tau}{\alpha_j}\|\boldsymbol{\theta}\|_1, \quad \text{subject to } \boldsymbol{W}\boldsymbol{\theta} \geq 0. \quad (1.10)$$

where $\boldsymbol{s}^j = \boldsymbol{\theta}^j - \frac{1}{\alpha_j}\nabla\phi(\boldsymbol{\theta}^j)$. Because the ℓ_1 norm is non-differentiable, a change of variables must be applied to (1.10) to use gradient-based methods. By letting $\boldsymbol{\theta} = \boldsymbol{u} - \boldsymbol{v}$ where $\boldsymbol{u}, \boldsymbol{v} \geq 0$, we can write (1.10) as

$$(\boldsymbol{u}^{j+1}, \boldsymbol{v}^{j+1}) = \arg\min_{(\boldsymbol{u},\boldsymbol{v})} \quad \frac{1}{2}\|\boldsymbol{s}^j - (\boldsymbol{u} - \boldsymbol{v})\|_2^2 + \frac{\tau}{\alpha_k}\mathbb{1}^t(\boldsymbol{u} + \boldsymbol{v}) \quad (1.11)$$

$$\text{subject to} \quad \boldsymbol{u} \geq 0, \ \boldsymbol{v} \geq 0, \ \boldsymbol{W}(\boldsymbol{u} - \boldsymbol{v}) \geq 0$$

where $\mathbb{1}$ is a vector of ones. The next iterate $\boldsymbol{\theta}^{j+1}$ is then defined as $\boldsymbol{\theta}^{j+1} = \boldsymbol{u}^{j+1} - \boldsymbol{v}^{j+1}$. Because the constraints on (1.11) are nonnegativity bounds not only on the variables $\boldsymbol{u}$ and $\boldsymbol{v}$ but also on $\boldsymbol{W}(\boldsymbol{u} - \boldsymbol{v})$, solving (1.11) is not straightforward. However, the *dual* formulation of this minimization problem, i.e., solving for the Lagrange multipliers associated with (1.11), has simple box constraints and can be easily solved iteratively. Using convex optimization theory, we can show that the primal solution, $\boldsymbol{\theta}^{j+1}$, can be obtained from the optimal Lagrange multipliers. We refer the reader to [29] for algorithmic details. In our numerical experiments below, we solve these subproblems approximately (i.e., with a limited number of iterations) as relatively good estimates of each primal iterate $\boldsymbol{\theta}^j$ tend to be produced with only few inner iterations.

1.4.3
Model-Based Sparsity

While the majority of the CS literature has focused on the case where the scene of interest admits a sparse representation in some basis or dictionary, more recent developments have used more sophisticated models of scenes which incorporate key *structure* into the sparsity models. The basic idea has been used previously in the context of image denoising and compression. For example, it is well-known that image denoising can be accomplished via wavelet coefficient-wise thresholding. However, more refined thresholding methods exploit the fact that significant wavelet coefficients tend to cluster near one another within scales and arise at similar locations between scales; this approach can yield significant improvements in accuracy [30].

CS reconstruction methods have recently been developed based upon similar principles to improve reconstruction results [31]. For instance, as we noted above, (1.10) amounts to an image denoising operation conducted during each loop of the reconstruction algorithm. Thus a variety of denoising methods can be used. Unfortunately, some of these models result in non-convex optimization problems for which finding the globally optimal reconstruction is not computationally tractable – though locally optimal reconstructions are often very good. Nevertheless, theoretical work in [31] shows that incorporating sparsity models can reduce the number of measurements needed to achieve a desired accuracy level. In the case where the image of interest is known to be smooth or piecewise smooth (i.e., it is compressible in a wavelet basis), we can formulate a penalty function which is a useful alternative to the ℓ_1 norm of the wavelet coefficient vector. In particular, we can build on the framework of *recursive dyadic partitions (RDP)*, which are described in detail in [32]. In particular, partition-based methods calculate image estimates by determining the ideal partition of the domain of observations and by fitting a model (e.g., a constant) to each cell in the optimal partition. This gives our estimators the capability of spatially varying the resolution to automatically increase the smoothing in very regular regions of the image and to preserve detailed structure in less regular regions.

An image $\boldsymbol{G}^{\mathrm{true}}$, an RDP $\boldsymbol{P}$, and an image approximation, $\boldsymbol{G}(\boldsymbol{P})$, are displayed in Fig. 1.5. RDP-based image estimates are computed using a very simple framework for penalized least squares estimation, wherein the penalization is based on the complexity of the underlying partition (i.e., the number of cells in the partition). The goal here is to find the partition which minimizes the penalized squared error function. Let $\boldsymbol{s}^j = \boldsymbol{g}^j - \frac{1}{\alpha_j}\nabla\phi(\boldsymbol{g}^j)$ where $\phi(\boldsymbol{g}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{g}\|_2^2$, and set

$$\boldsymbol{P}^{j+1} = \underset{\boldsymbol{P}\in\mathcal{P}}{\arg\min}\ \|\boldsymbol{g}(\boldsymbol{P}) - \boldsymbol{s}^j\|_2^2 + \mathrm{pen}(\boldsymbol{P})\ \text{ subject to } \boldsymbol{g}(\boldsymbol{P}) \geq 0$$
$$\boldsymbol{g}^{j+1} = \boldsymbol{g}(\boldsymbol{P}^{j+1})$$

where $\mathcal{P}$ is a collection of all possible RDPs of the image domain and $\boldsymbol{g}(\boldsymbol{P}) = \mathrm{vect}(\boldsymbol{G}(\boldsymbol{P}))$. A search over partitions can be computed quickly using a dynamic program. The model coefficients for each partition cell are chosen via nonnegative least squares. enforcing nonnegativity constraints is trivial and can be accomplished very quickly.
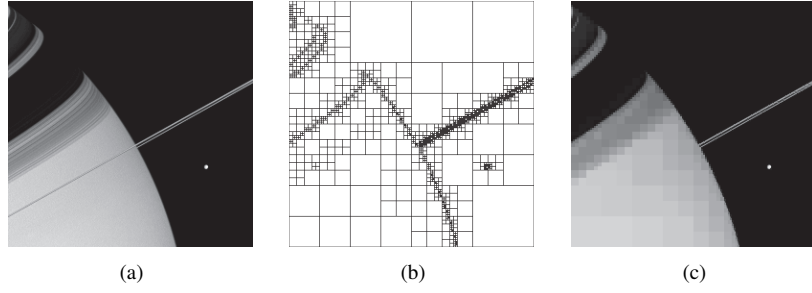
(a) (b) (c)

Fig. 1.5 Recursive dyadic partition of an image. (a) Original image G^{true}. (b) RDP P, with larger partition cells corresponding to regions of more homogeneous intensity. (c) Piecewise constant approximation $G(P)$ to original image, with constant pieces corresponding to RDP cells.

1.4.4

Compensating for Nonnegative Sensing Matrices

Generative models for random projection matrices used in CS often involve drawing elements independently from a zero-mean probability distribution [1, 3, 21], and likewise a zero-mean distribution is used to generate the coded aperture masks described in Sec.1.3.1. However, a coded aperture mask with a zero mean is not physically realizable in optical systems. We generate our physically realizable mask by taking our randomly generated, zero-mean mask pattern and shifting it so that all mask elements are in the range $[0, 1/m]$, where m is the dimension of the observation [12]. This shifting ensures that the coded aperture corresponds to a valid (i.e., nonnegative and intensity preserving) probability transition matrix which describes the distribution of photon propagation through the optical system.

This shifting, while necessary to accurately model real-world optical systems, negatively impacts the performance of the proposed ℓ_2-ℓ_1 reconstruction algorithm for the following reason. If we generate a non-realizable zero-mean mask (H^0) with elements in the range $[-1/2m, 1/2m]$ and simulate measurements of the form

$$y^0 = A^0 g^{\text{true}} \triangleq \text{vect}(\mathcal{D}(G^{\text{true}} * H^0)), \qquad (1.12)$$

then the corresponding observation matrix A^0 will satisfy the RIP with high probability and g^{true} can be accurately estimated from y^0 using ℓ_2-ℓ_1 minimization. In contrast, if we rescale H^0 to be in the range $[0, 1/m]$ and denote this H, then we have a practical and realizable coded aperture mask. However, observations of the form

$$y = A g^{\text{true}} \triangleq \text{vect}(\mathcal{D}(G^{\text{true}} * H))$$

lead to an objective function whose second derivative is no longer accurately approximated by a scalar multiple of the identity matrix, thereby mitigating the effectiveness of methods that exploit this property. This problem is addressed in [12], where it is shown that the zero-mean observation y^0 can be estimated by shifting each measurement in y. Let C_A is the sum of each column of A and $\mu \triangleq \sum_{i=1}^{m} y_i / C_A$; note

$\boldsymbol{y}^0 \approx \boldsymbol{y} - (\mu/2m)\mathbb{1}_{m \times 1}$. We can then solve

$$\boldsymbol{\theta}^{\mathbf{est}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2}\|\boldsymbol{y}^0 - \boldsymbol{A}^0\boldsymbol{W}\boldsymbol{\theta}\|_2^2 + \mathrm{pen}(\boldsymbol{\theta}) \quad \text{subject to} \quad \boldsymbol{W}\boldsymbol{\theta} + \mu\mathbb{1} \geq 0 \quad (1.13)$$

where $\mathrm{pen}(\boldsymbol{\theta})$ is either the ℓ_1 or the model-based sparsity penalty. The estimate for $\boldsymbol{g}^{\mathrm{est}}$ is then given by $\boldsymbol{g}^{\mathrm{est}} = \boldsymbol{W}\boldsymbol{\theta}^{\mathrm{est}} + \mu\mathbb{1}$. We note that the algorithms described in Sec. 1.4.2 and 1.4.3 can be modified to solve (1.13) in a straightforward manner.

1.5

Experimental Results

The utility of optical system designs based on CS theory for improving the resolution of optical systems is demonstrated via a simulation study in this section. In particular, we consider several different optical mechanisms which could be used to image Saturn and its rings using a "ground truth" image displayed in Fig. 1.5(a); this image was originally obtained by the Cassini Orbiter [33] and was cropped and downsampled to size 256×256 for our numerical experiments. The pixel intensities range from 0 to 4.53. We add white Gaussian noise to the projected measurements $\boldsymbol{A}\boldsymbol{g}^{\mathrm{true}}$ to model the sensor noise associated with the focal plane array. In the pinhole and coded aperture experiments, we acquire m measurements over a total time T sec. We assume that this averaged noise has standard deviation $\sigma = 10^{-5}$. The single-pixel camera collects measurements sequentially, and consequently, it has at most T/m sec to obtain each of the m measurements. Thus, the noise variance associated with the single-pixel camera model must be scaled appropriately, though this disadvantage is offset by the fact that the single-pixel camera can implement arbitrary random projects and hence satisfy the RIP for smaller FPA dimension m.

Fig. 1.6(a) shows the result of observing the scene with a pinhole camera and an FPA of size 128×128. This image is clearly lower in resolution than the original (i.e., it is an image with 128×128 pixels, upsampled here via nearest neighbor interpolation for visualization and computing errors). It also exhibits noise artifacts due to the relatively low light intensity reaching the camera detectors relative to the noise level. The root-mean-squared (RMS) error of this image, $\|\boldsymbol{g}^{\mathrm{est}} - \boldsymbol{g}^{\mathrm{true}}\|_2/\|\boldsymbol{g}^{\mathrm{true}}\|_2$, is 0.117. Conventional coded aperture techniques, such as the MURA methods described above, can be used to improve the noise performance of a camera. This is demonstrated in Fig. 1.6(b), where the MURA code has a fill factor of 50% and the reconstruction has an RMS error of 0.071. However, these conventional coded aperture systems cannot resolve details smaller than the pixel size of the FPA. Note the pixelization of Saturn's moon (Mimas), rings, and edges in the MURA reconstruction.

Next, we describe the improvements gained by taking compressive measurements. First, we simulate the image acquisition process in the single-pixel camera by taking random projections of the image and then solve the corresponding CS minimization problem. The reconstruction for the single-pixel camera is shown in Fig. 1.6(c). We note the improved resolution of Mimas and Saturn's rings. For brevity, we only show

the reconstruction from the RDP-based method we proposed, which significantly out-performed the other methods we discussed.

We also observe that the compressive coded aperture techniques described in this chapter can lead to further gains. For instance, Fig. 1.6(d) shows the reconstructed 256×256 image computed via GPSR [24] (using the Haar wavelet basis) from CCA data collected on a 128×128 FPA with a 50% fill factor coding mask; the RMS error here is 0.067. (While a variety of different convergence criteria have been successfully employed for the reconstruction algorithms, in this section we simply display the result after 100 iterations. A more detailed perspective on RMS error decay across iterations or computation time is described in [29]; that work considers a Poisson noise model but otherwise equivalent reconstruction methods to the ones considered here.) We note that this result has over 4200 negatively-valued pixels. Additional gains are possible by incorporating positivity constraints, as shown in Fig. 1.6(e), where the RMS error is 0.049 and none of the pixels have negative values.

Finally, including structure in sparsity models during the reconstruction process via RDPs yields additional improvements. Using RDP-based denoising within our optimization framework yields an RMS error of 0.041 and zero pixels with negative values. However, a cycle-spun version of RDP-based denoising further improves performance, as shown in Fig. 1.6(f), where the RMS error is 0.032, a 73% improvement over the pinhole camera result, 55% improvement over the MURA result, and a 50% improvement over the single-pixel camera result. In addition, Fig. 1.6 shows that compressive coded apertures significantly improve our ability to resolve fine features, such as the side-view of Saturn's rings, its moon, and bands in the planet's atmosphere. We note that the ℓ_2-ℓ_1 CS minimization with nonnegativity constraints and the partition-based approaches (Figs. 1.6(e) and (f)) were initialized using the GPSR result thresholded to satisfy the nonnegativity constraints.

One other advantage that compressive coded apertures have over the Rice single-pixel camera is the computational time necessary to perform reconstruction. Because the observation matrices associated with CCA involve Fourier transforms and diagonal matrices, matrix-vector products are very fast and memory efficient. In contrast, the numerical methods associated with the Rice single-pixel camera are significantly slower and more memory intensive due to the dense, unstructured projection matrix.

1.6

Noise and Quantization

While CS is particularly useful when the FPA needs to be kept compact, it should be noted that CS is more sensitive to measurement errors and noise than more direct imaging techniques. The experiments conducted in this chapter simulated very high signal-to-noise ratio (SNR) settings and showed that CS methods can help resolve high resolution features in images. However, in low SNR settings CS reconstructions can exhibit significant artifacts that may even cause more distortion than the low-resolution artifacts associated with conventional coded aperture techniques such as MURA.
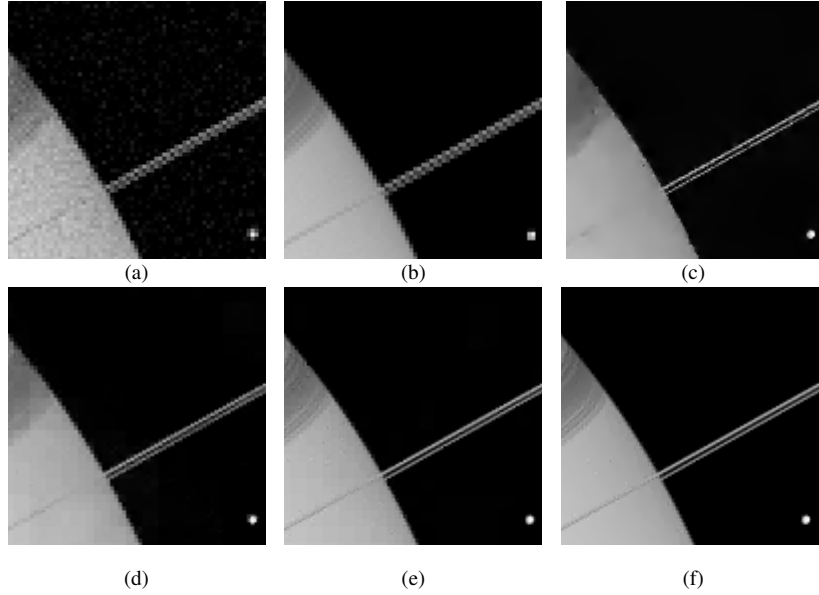
Fig. 1.6 Simulation results comparing different image acquisition and reconstruction strategies. (a) Observations from a pinhole aperture, upsampled using nearest-neighbor interpolation; RMS error = 0.117. (b) Reconstruction from a MURA coded aperture camera system, upsampled using nearest-neighbor interpolation; RMS error = 0.071. (c) Translationally invariant RDP-base reconstruction from a single-pixel camera system; RMS error = 0.064. (d) ℓ_2-ℓ_1 reconstruction from a CCA camera system; RMS error = 0.067. (e) ℓ_2-ℓ_1 (with nonnegativity constraint) reconstruction from a CCA Camera system; RMS error = 0.049. (f) Translationally invariant RDP-based reconstruction; RMS error = 0.032. Note the higher resolution in the reconstructions (d)-(f) from a CCA camera system and how performance can be further improved using cycle-spun RDP-based reconstruction methods.

The coded aperture work discussed earlier is focused on developing a coded aperture system with a *fill factor* of 50%; i.e., 50% of the positions in the coded mask were opaque, and the remainder allowed light through to the detector. In low-light settings, this high fill factor is desirable because it allows a significant proportion of the light through to the detector and "wastes" very few photons. This approach is particularly effective when the scene is sparse in the canonical or pixel basis (e.g., faint stars against a dark sky). However, when the scene is sparse in some other basis, such as a wavelet basis, and *not* sparse in the pixel basis, a large fill factor can cause significant noise challenges. These challenges are described from a theoretical perspective in [28]. Intuitively, consider that when the fill factor is close to 50%, most of the coded aperture measurements will have the same average intensity plus a small fluctuation, and the photon noise level will scale with this average intensity. As a result, in low light settings the noise will overwhelm the small fluctuations which are critical to accurate reconstruction unless the scene is sparse in the pixel basis and the average intensity per pixel is low. These challenges can be mitigated somewhat by using smaller fill

factors, but in general limit the utility of *any* linear optical CS architecture for very low-intensity images which are not sparse in the canonical basis [28].

Similar observations are made in [34], which presents a direct comparison of the noise robustness of CS in contrast to conventional imaging techniques both in terms of bounds on how reconstruction error decays with the number of measurements and in a simulation setup; the authors conclude that for most real-world images, CS yields the biggest gains in high signal-to-noise ratio (SNR) settings. These observations are particularly relevant when considering the bit-depth of focal plane arrays, which corresponds to measurement quantization errors. Future efforts in designing optical CS systems must carefully consider the amount of noise anticipated in the measurements to find the optimal tradeoff between the focal plane array size and image quality. (For a description of image quality assessment techniques see Ch. 19.)

1.7
Conclusions

This chapter describes several recent efforts aimed at the exploitation of important theoretical results from compressive sensing (CS) in practical optical imaging systems. One of the main tenets of CS is that relatively few well-chosen observations can be used to form a sparse image using sophisticated image reconstruction algorithms. This suggests that it may be possible to build cameras with much smaller focal plane arrays than are conventionally required for high-resolution imaging. However, the application of these ideas in practical settings poses several challenges.

First, directly implementing CS theory by collecting a series of independent pseudo-random projections of a scene requires either (a) a very large physical system or (b) observations collected sequentially over time. This latter approach is successfully used, for instance, in the Rice Single Pixel Camera. Alternative snapshot architectures (which capture all observations simultaneously) with a compact form factor include coded aperture techniques. These approaches impose structure upon the pseudo-random projections, most notably by limiting their independence. As a result, the number of measurements required to accurately recover an image is higher with snapshot coded aperture systems.

A second key challenge relates to the nonnegativity of image intensities and measurements which can be collected by linear optical systems. Much of the theoretical literature on CS allows for negative measurements and does not consider nonnegativity during the reconstruction process. In this chapter we have shown that (a) explicitly incorporating nonnegativity constraints can improve reconstruction accuracy and (b) pre-processing observations to account for nonnegative sensing matrices improves reconstruction performance because of central assumptions underlying some fast CS algorithms. However, one important open question is whether novel approaches based on *nonlinear* optics can successfully circumvent these positivity constraints to improve performance.

References

1 Candès, E. J. and Tao, T., "Decoding by linear programming," *IEEE Trans. Inform. Theory*, 15, pp. 4203–4215, 2005.

2 Donoho, D. L., "Compressed sensing," *IEEE Trans. Inform. Theory*, 52, pp. 1289–1306, 2006.

3 Haupt, J. and Nowak, R., "Signal reconstruction from noisy random projections," *IEEE Trans. Inform. Theory*, 52, pp. 4036–4048, 2006.

4 Natarajan, B. K., "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, 24, pp. 227–234, 1995.

5 Tibshirani, R., "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, 58, pp. 267–288, 1996.

6 Tropp, J. A., "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inform. Theory*, 52, pp. 1030–1051, 2006.

7 Neifeld, M. and Ke, J., "Optical architectures for compressive imaging," *Appl. Opt.*, 46, pp. 5293–5303, 2007.

8 Duarte, M. F. et al., "Single pixel imaging via compressive sampling," *IEEE Sig. Proc. Mag.*, 25, pp. 83–91, 2008.

9 Gehm, M. et al., "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express*, 15, pp. 14013–14027, 2007.

10 Maggioni, M. et al., "Hyperspectral microscopic analysis of normal, benign and carcinoma microarray tissue sections," in "Proc. of SPIE, Optical Biopsy VI," volume 6091, 2006.

11 Ashok, A., Baheti, P. K. and Neifeld, M. A., "Compressive imaging system design using task-specific information," *Appl. Opt.*, 47, pp. 4457–4471, 2008.

12 Marcia, R. F. and Willett, R. M., "Compressive coded aperture superresolution image reconstruction," in " Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing," 2008.

13 Stern, A. and Javidi, B., "Random projections imaging with extended space-bandwidth product," *IEEE/OSA J. Disp. Technol.*, 3, pp. 315–320, 2007.

14 Dicke, R. H., "Scatter-hole cameras for X-rays and gamma-rays," *Astrophysical Journal*, 153, pp. L101–L106, 1968.

15 Skinner, G., "Imaging with coded-aperture masks," *Nucl. Instrum. Methods Phys. Res.*, 221, pp. 33–40, 1984.

16 Accorsi, R., Gasparini, F. and Lanza, R. C., "A coded aperture for high-resolution nuclear medicine planarimaging with a conventional anger camera: experimental results," *IEEE Trans. Nucl. Sci.*, 28, pp. 2411–2417, 2001.

17 Levin, A. et al., "Image and depth from a conventional camera with a coded aperture," in "Proc. of Intl. Conf. Comp. Graphics. and Inter. Tech.", 2007.

18 Sloane, N. J. A. and Harwit, M., "Masks for Hadamard transform optics, and

weighing designs," *Appl. Opt.*, 15, p. 107, 1976.

19 Ashok, A. and Neifeld, M. A., "Pseudo-random phase masks for superresolution imaging from subpixel shifting," *Appl. Opt.*, 46, pp. 2256–2268, 2007.

20 Gottesman, S. R. and Fenimore, E. E., "New family of binary arrays for coded aperture imaging," *Appl. Opt.*, 28, 1989.

21 Bajwa, W. et al., "Toeplitz-structured compressed sensing matrices," in "Proc. of IEEE Stat. Sig. Proc. Workshop," 2007.

22 Romberg, J., "Compressive sampling by random convolution," 2009, to appear in *SIAM J. Sci. Comput.*

23 Kim, S. J. et al., "An interior-point method for large-scale ℓ_1-regularized least squares," *IEEE J. Sel. Top. Sign. Proces.*, pp. 606–617, 2007.

24 Figueiredo, M. A. T., Nowak, R. D. and Wright, S. J., "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Sign. Proces.: Special Issue on Convex Optimization Methods for Signal Processing*, 1, pp. 586–597, 2007.

25 Wright, S., Nowak, R. and Figueiredo, M., "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, 2009 (to appear).

26 Yin, W. et al., "Bregman iterative algorithms for l1-minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, 1, pp. 143–168, 2008.

27 Tropp, J. A., "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, 50, pp. 2231–2242, 2004.

28 Raginsky, M. et al., "Compressed sensing performance bounds under Poisson noise," 2009, submitted.

29 Harmany, Z., Marcia, R. and Willett, R., "Sparse Poisson intensity reconstruction algorithms," in "Proc. IEEE Stat. Sig. Proc. Workshop," 2009.

30 Crouse, M. S., Nowak, R. D. and Baraniuk, R. G., "Wavelet-based statistical signal-processing using hidden markov-models," *IEEE Trans. Sig. Proc.*, 46, pp. 886–902, 1998.

31 Baraniuk, R. et al., "Model-based compressive sensing," *IEEE Trans. on Inform. Theory*, 2008, submitted.

32 Willett, R. and Nowak, R., "Multiscale Poisson intensity and density estimation," *IEEE Trans. Inform. Theory*, 53, pp. 3171–3187, 2007.

33 NASA/JPL/Space Science Institute, "Pia10497: Saturn in recline," 2008, http://photojournal.jpl.nasa.gov/catalog/PIA10497.

34 Haupt, J. and Nowak, R., "Compressive sampling vs. conventional imaging," in "Proc. IEEE Intl. Conf. on Imaging Processing," pp. 1269–1272, 2006.