

Block compressed sensing of images and video

By

Sungkwang Mun

A Dissertation  
Submitted to the Faculty of  
Mississippi State University  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy  
in Computer Engineering  
in the Department of Electrical and Computer Engineering

Mississippi State, Mississippi

December 2012

Copyright by  
Sungkwang Mun  
2012

Block compressed sensing of images and video

By

Sungkwang Mun

Approved:

---

James Fowler  
Professor of Electrical and Computer  
Engineering, and Graduate Coordinator  
(Major Professor)

---

Lori Bruce  
Professor of Electrical and Computer  
Engineering and Associate Dean of  
Research (Committee Member)

---

Qian Du  
Associate Professor of Electrical and  
Computer Engineering  
(Committee Member)

---

Song Zhang  
Associate Professor of Computer  
Science and Engineering  
(Committee Member)

---

Sarah A. Rajala  
Dean of the James Worth Bagley College  
of Engineering

Name: Sungkwang Mun

Date of Degree: December 15, 2012

Institution: Mississippi State University

Major Field: Computer Engineering

Major Professor: Dr. James Fowler

Title of Study: Block compressed sensing of images and video

Pages of Study: 143

Candidate for Degree of Doctor of Philosophy

Compressed sensing is an emerging approach for signal acquisition wherein theory has shown that a small number of linear, random projection of a signal contains enough information for reconstruction of the signal. Despite its potential to enable lightweight and inexpensive sensing hardware that simultaneously combines signal acquisition and dimensionality reduction, the compressed sensing of images and video still entails several challenges, in particular, a sensing-measurement operator which is difficult to apply in practice due to the heavy memory and computational burdens. Block-based random image sampling coupled with a projection-driven compressed-sensing recovery is proposed to address this challenge.

For images, the block-based image acquisition is coupled with reconstruction driven by a directional transform and statistical model based thresholding that encourages spatial sparsity. Also considered is an extension of the basic reconstruction algorithm that incorporates block-based measurements in the domain of a wavelet transform. The pro-

posed image recovery algorithm and its extension yield images with quality that matches or exceeds that produced by a popular, yet computationally expensive, technique which minimizes total variation with a significantly less computational complexity.

For video, motion estimation and compensation is utilized to promote temporal sparsity. A residual between the current frame and the previous frame compensated by object motion is shown to be more sparse than the original frame itself. By using residual reconstruction, information contained in the previous frame contributes to the reconstruction of the current frame. The proposed block-based compressed-sensing reconstruction for video outperforms a simple frame-by-frame reconstruction as well as a 3D volumetric reconstruction in terms of visual quality.

Finally, quantization of block-based compressed-sensing measurements is considered in order to generate a true bitstream from a compressed-sensing image acquisition. Specifically, a straightforward process of quantization via simple uniform scalar quantization applied in conjunction with differential pulse code modulation of the block-based compressed-sensing measurements is proposed. Experimental results demonstrate significant improvement in rate-distortion performance as compared scalar quantization used alone in several block-based compressed-sensing reconstruction algorithms as well as that of alternative quantized-compressed-sensing techniques relying on optimized quantization or reconstruction is observed.

## DEDICATION

To my wife, NaYeon, who support me in all ways through the long hours of research  
and writing

## ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. James E. Fowler who has not only advised me but also been a good role model as a professor. Without his encouragement and guidance, this work could not have been achieved. I also thank to Dr. Lori Bruce, Dr. Qian Du, and Dr. Song Zhang for working as committee members. I am also thankful to my good friend and colleague Eric Tramel at MSU for his generous help and helpful discussion. This work has been supported by NSF whose generous funding I greatly appreciate. My deepest gratitude goes to my parents whose immeasurable support and love have brought me here. Finally, I would like to give this honor to God who created me.

## TABLE OF CONTENTS

DEDICATION . . . . .	ii
ACKNOWLEDGEMENTS . . . . .	iii
LIST OF TABLES . . . . .	vi
LIST OF FIGURES . . . . .	vii
CHAPTER	
1. INTRODUCTION . . . . .	1
2. BLOCK-BASED COMPRESSED SENSING FOR STILL IMAGES . . .	5
2.1 An Overview of Compressed Sensing Theory . . . . .	6
2.1.1 Approaches to CS-Based Signal Acquisition . . . . .	9
2.1.2 Approaches to CS Reconstruction . . . . .	11
2.2 CS Acquisition of Still Images . . . . .	15
2.2.1 The Single-Pixel Camera . . . . .	15
2.2.2 Other Image-Acquisition Architectures . . . . .	19
2.3 Straightforward Reconstruction for Images . . . . .	20
2.4 Total-Variation Reconstruction . . . . .	23
2.5 CS with Blocks in the Spatial Domain . . . . .	24
2.5.1 The BCS-SPL Algorithm . . . . .	27
2.5.2 BCS-SPL with Directional Transforms . . . . .	29
2.5.3 Experimental Observations on BCS-Based Reconstruction	31
2.6 CS with Blocks in the Wavelet-Domain . . . . .	37
2.6.1 Wavelet-Domain BCS . . . . .	38
2.6.2 Wavelet-Domain MS-BCS-SPL . . . . .	41
2.7 Other Approaches to CS Reconstruction of Images . . . . .	43
2.8 Comparison of Various CS Techniques for Images . . . . .	45
2.9 Remarks . . . . .	50
3. BLOCK-BASED COMPRESSED SENSING FOR VIDEO . . . . .	53



3.1	Block-Based Acquisition of Video . . . . .	54
3.2	Straightforward CS Reconstruction for Video . . . . .	55
3.3	The Motion-Compensated BCS-SPL Algorithm . . . . .	56
3.3.1	Residual Reconstruction . . . . .	57
3.3.2	Multihypothesis Initialization . . . . .	61
3.3.3	Forward/Backward Motion Compensation . . . . .	62
3.4	Experimental Observations . . . . .	68
3.4.1	MC-BCS-SPL with Equal Subrate . . . . .	71
3.4.2	MC-BCS-SPL with Key Frames of Increased Subrate . . .	71
3.5	Comparison of Various CS Techniques for Video . . . . .	80
3.6	Remarks . . . . .	85
4.	DPCM FOR QUANTIZED BLOCK-BASED COMPRESSED SENSING OF IMAGES . . . . .	87
4.1	Background . . . . .	89
4.1.1	Quantized Compressed Sensing . . . . .	89
4.1.2	DPCM for Natural Images . . . . .	91
4.1.3	Correlation of Measurements in BCS . . . . .	95
4.2	DPCM for Quantized BCS-SPL . . . . .	98
4.2.1	1D-DPCM for Quantized BCS-SPL . . . . .	99
4.2.2	Coding Gain of Quantized BCS-SPL . . . . .	102
4.2.3	2D-DPCM for Quantized BCS-SPL . . . . .	107
4.3	Experimental Observation . . . . .	108
4.3.1	1D-BCS-SPL-DPCM with SQ . . . . .	110
4.3.2	2D-BCS-SPL-DPCM with SQ . . . . .	112
4.3.3	Comparison of Various CS Techniques for Quantized Mea- surements . . . . .	123
4.4	Remarks . . . . .	130
5.	CONCLUSIONS . . . . .	131

## LIST OF TABLES

2.1	BCS-based reconstruction performance as measured in PSNR in dB . . . .	36
2.2	Wavelet-domain BCS subrates $S_l$ at level $l$ for target overall subrate $S$ . . .	40
2.3	Reconstruction PSNR in dB; images are $512 \times 512$ . . . . .	47
2.4	Reconstruction time for the $512 \times 512$ “Lenna” image . . . . .	48
3.1	Two-phase multiframe reconstruction for a GOP of size $P = 8$ . . . . .	68
3.2	Average PSNR in dB for several video sequences; $S_K = 0.7$ , 88 frames . .	70
3.3	Average PSNR in dB for several video sequences; $S_K = 0.7$ , 88 frames . .	82
3.4	Reconstruction time in seconds per frame (spf) . . . . .	85
4.1	PSNR Performance in dB for a bitrate of 0.5 bpp . . . . .	110

## LIST OF FIGURES

2.1	The single-pixel camera for the CS acquisition of a still image . . . . .	17
2.2	BCS-SPL reconstruction of a 2D image . . . . .	28
2.3	Reconstructions of the $512 \times 512$ “Lenna” image . . . . .	34
2.4	Reconstructions of the $512 \times 512$ “Lenna” image . . . . .	35
2.5	MS-BCS-SPL reconstruction of a 2D image . . . . .	42
2.6	Reconstructions of the $512 \times 512$ “Lenna” image . . . . .	49
3.1	Decay of the magnitudes of the transform coefficients . . . . .	59
3.2	MC-BCS-SPL reconstruction from a single reference frame . . . . .	60
3.3	Multihypothesis initialization in MC-BCS-SPL reconstruction . . . . .	62
3.4	Various strategies for the initial estimate of the current frame . . . . .	63
3.5	MC-BCS-SPL reconstruction applied to a GOP of size $P = 8$ frames . . .	67
3.6	Performance of MC-BCS-SPL on the “Susie” for $S_K = S_{NK}$ . . . . .	72
3.7	Performance of MC-BCS-SPL on the “Football” for $S_K = S_{NK}$ . . . . .	73
3.8	Performance of MC-BCS-SPL on the “Mother and Daughter” for $S_K \neq S_{NK}$	76
3.9	Performance of MC-BCS-SPL on the “Hall Monitor” for $S_K \neq S_{NK}$ . . . .	77
3.10	Reconstructed center frame for “Football” . . . . .	78
3.11	Reconstructed center frame for “Susie” . . . . .	79
3.12	Reconstructions of frame 4 of the “Foreman” . . . . .	83

3.13	Reconstructions of frame 4 of the “Mother and Daughter” . . . . .	84
4.1	DPCM for an image . . . . .	93
4.2	Correlation of measurements in CS and BCS . . . . .	97
4.3	Application of DPCM and SQ to the BCS-SPL architecture . . . . .	99
4.4	Error-bar plot with respect to block sizes . . . . .	103
4.5	Plot of RD efficiency (PSNR/rate) on various block sizes . . . . .	104
4.6	Prediction gain and average of normalized correlation coefficients . . . . .	106
4.7	Application of 2D-DPCM and SQ to the BCS-SPL architecture . . . . .	108
4.8	The convex rate-distortion hull . . . . .	111
4.9	RD performance of 1D-DPCM plus SQ applied to BCS-SPL . . . . .	113
4.10	RD performance of 1D-DPCM plus SQ applied to MS-BCS-SPL . . . . .	114
4.11	RD performance of 1D-DPCM plus SQ applied to MH-BCS-SPL . . . . .	115
4.12	RD performance of various predictions in 2D-DPCM for Lenna . . . . .	116
4.13	RD performance of various predictions in 2D-DPCM for Barbara . . . . .	117
4.14	RD performance of various predictions in 2D-DPCM for Peppers . . . . .	118
4.15	RD performance of various predictions in 2D-DPCM for Goldhill . . . . .	119
4.16	RD performance of 2D-DPCM plus SQ applied to BCS-SPL . . . . .	120
4.17	RD performance of 2D-DPCM plus SQ applied to MS-BCS-SPL . . . . .	121
4.18	RD performance of 2D-DPCM plus SQ applied to MH-BCS-SPL . . . . .	122
4.19	Rate-distortion performance for Lenna . . . . .	124
4.20	Rate-distortion performance for Barbara . . . . .	125
4.21	Rate-distortion performance for Peppers . . . . .	126

4.22	Rate-distortion performance for Goldhill . . . . .	127
4.23	Reconstructions of the $512 \times 512$ “Lenna” image . . . . .	129

## CHAPTER 1

### INTRODUCTION

The *sampling theorem* is arguably the best known component of the theoretical foundations of the signal-processing and communications fields; its importance is paramount in that it underlies all modern signal-acquisition, sampling, sensing, and analog-to-digital conversion devices. Although introduced to the signal-processing and communications communities by Shannon in 1949 [92], the sampling theorem can be traced to earlier work by telegraphers and mathematicians (see, e.g., [75]). In essence, the sampling theorem states that, if a signal contains no frequencies higher than bandlimit  $W$ , then it can be perfectly reconstructed from samples acquired at a rate of at least  $2W$ . This latter quantity, commonly known as the *Nyquist rate*, thus represents the slowest rate at which sampling of any bandlimited signal can be acquired and still permit perfect reconstruction.

However, this traditional sampling theory is founded on relatively minimal prior knowledge on the signal being sampled—i.e., its bandlimit  $W$ . While traditional sampling theory has the advantage of applying to *any* signal satisfying this bandlimit constraint, we are commonly interested in more restricted classes of signals, i.e., those that are known to possess much more structure, and thus fewer degrees of freedom, than dictated by the signal bandlimit [7]. A well-known example is that of bandpass signals in which the signal is present over only a limited band of frequencies—under such bandpass structure, bandpass

sampling (e.g., [105]) can acquire the signal with a sampling rate slower than  $2W$ . More recent literature has witnessed an explosion of interest in sensing that exploits structured prior knowledge in the general form of *sparsity*, meaning that signals can be represented by only a few coefficients in some transform basis. Like bandpass sampling, exploitation of such sparse structure within signals can effectively permit sampling at rates well below  $2W$ .

Central to much of this recent work is the paradigm of compressed sensing (CS)<sup>1</sup> (e.g., [18, 22, 37]) which permits relatively few measurements of the signal to be acquired in a linear fashion while still permitting exact reconstruction via a relatively complex and non-linear recovery process. While much CS literature is rather generic in that it is not tied to any specific class of signal beyond a general assumption of sparsity, there has been significant interest in CS specifically tailored to imaging applications. Indeed, recent work in the CS field has seen proposals for not only sensor devices but also reconstruction algorithms designed specifically for a variety of imagery signals.

In this dissertation, we focus on photographic imagery which is acquired in the spatial domain of the image, a paradigm which is ubiquitous throughout image-processing applications. This stands in contrast to a significant portion of existing CS literature that has targeted, with substantial success, specific medical-imaging applications—in particular, magnetic resonance imaging (MRI), which is acquired directly in a Fourier-transform space. The potential for CS to significantly expedite MRI acquisition is relatively well established and already well covered tutorially in the literature (e.g., [76, 77]). On the

---

<sup>1</sup>also known as *compressive sampling* or *compressive sensing*

other hand, CS for photographic imagery acquired in the spatial domain is a comparatively emerging area and, thus, the topic of the present dissertation.

A primary contribution of this dissertation comprises the development of block-based CS reconstruction. In such block-based compressed sensing (BCS), an image is partitioned into small non-overlapping blocks which are acquired independently but reconstructed jointly. BCS is motivated primarily for reasons of reduced computational complexity and memory burdens which can become impractically large for the CS of images and video as a result of the increased dimensionality (i.e., 2D and 3D) of such signals. We examine the performance of the BCS recovery methodology by comparing the state of the art in CS reconstruction for various imaging modalities, including still images and motion video.

Additionally, while much of CS literature focuses on the CS signal-acquisition process as merely a form of dimensionality reduction (i.e., a measurement process wherein the number of samples obtained is much less than the dimensionality of the original signal), practical applications of CS will require that some form of quantization be applied in order for a CS sensor to generate a true bitstream. An additional contribution of this dissertation is the investigation of a quantized CS scheme wherein simple uniform scalar quantization (SQ) is coupled with differential pulse-code modulation (DPCM) of BCS measurements. We find through a battery of experimental results that such DPCM-based quantization yields rate-distortion performance significantly superior to that of the direct uniform SQ of CS measurements as well as that of several other quantized-CS schemes from recent literature.



We note that the material that comprises this dissertation has previously appeared in several publications. Specifically, our work on BCS was introduced originally for still images in [48, 82] and for video in [83]; this work constitutes Chapters 2 and 3, respectively, of this dissertation. Our work on DPCM-quantized CS appeared initially in [84] and forms Chapter 4 here. Finally, a comprehensive and in-depth survey of all our work on BCS for images and video was published as [49].

The remainder of this dissertation is organized as follows. Chapter 2 briefly overviews CS theory, including acquisition and reconstruction processes, and then considers the BCS of a single still image, comparing with a variety of techniques to reconstruct such images from random BCS measurements. Chapter 3 extends these concepts to the BCS of video with an emphasis on reconstruction from motion-compensated residuals. Chapter 4 finally considers quantized BCS measurements wherein rate-distortion performance is significantly improved by coupling DPCM with SQ. Finally, we bring the discussion to a close by making several concluding remarks in Chapter 5.

## CHAPTER 2

### BLOCK-BASED COMPRESSED SENSING FOR STILL IMAGES

Compressed sensing (CS) is a framework for acquiring and recovering a sparse signal from samples, or measurements, taken at rate significantly below that entailed by traditional sampling theory. Unlike traditional point-by-point sampling, the CS acquisition process involves random projection of the signal. In other words, the CS measurement process is a linear projection into a lower-dimensional subspace chosen at random. Theoretical results shows that it is possible to find an exact solution during reconstruction if the signal is exactly sparse in some transform domain. The theory also applies to the approximate reconstruction of signals that are only approximately sparse. Because images and video signals are usually approximately sparse in some transform domain such as Fourier or wavelet, CS can successfully recover the such image signals.

One of the most significant potential benefits of CS is its promise of providing simultaneous sampling and dimensionality reduction. Due to randomness of the CS projection, each measurement carries roughly the same amount of information, which means that all measurements are equally important [31]. Therefore, it is possible to reduce the dimension without sorting out the measurements that preserve significant information as is often done in traditional compression systems. This property promises a simpler encoding system with much fewer samples acquired at the sensor.

However, there are some challenges to practical CS imaging systems. The random projection patterns, which are usually represented as a matrix, must be stored at both the acquisition and recovery sides of the CS system, and the size of the measurement matrix increases quickly as the size of signal of interest increases. Moreover, a CS measurement matrix with a huge size can cause significant computation during reconstruction if explicit matrix operations are involved. In this chapter, we propose a block-based methodology to substantially mitigate these issues.

In this chapter, we give a brief overview of the CS theory in Section 2.1 and then describe an image-acquisition framework for CS in Section 2.2. Sections 2.3 and 2.4 illustrate the general image reconstruction process as well as a specific algorithm which finds a sparse solution in the gradient domain, respectively. Our proposed block-based algorithm, block-based compressed sensing with smooth projected Landweber reconstruction (BCS-SPL), is described in Section 2.5, and its extension—multiscale BCS-SPL (MS-BCS-SPL)—is introduced in Section 2.6. After detailing other approaches for CS image recovery in Section 2.7, reconstruction performance is compared experimentally in Section 2.8. We make several concluding remarks for the chapter in Section 2.9. We note that much of the material of this chapter appeared in print previously as [48, 49, 82].

## **2.1 An Overview of Compressed Sensing Theory**

In brief, CS [18, 22, 37] is an emerging mathematical paradigm which permits, under certain conditions, linear projection of a signal into a dimension much lower than that of the original signal while allowing exact recovery of the signal from the projections. More

specifically, suppose that we want to recover real-valued signal  $\mathbf{x}$  with length  $N$  from  $M$  measurements such that  $M \ll N$ . In other words, we want to recover  $\mathbf{x} \in \mathbb{R}^N$  from

$$\mathbf{y} = \Phi \mathbf{x}, \quad (2.1)$$

where  $\mathbf{y}$  has length  $M$ , and  $\Phi$  is an  $M \times N$  measurement matrix with subsampling rate, or *subrate*, being  $S = M/N$ . Because the number of unknowns is much larger than the number of observations, recovering every  $\mathbf{x} \in \mathbb{R}^N$  from its corresponding  $\mathbf{y} \in \mathbb{R}^M$  is impossible in general; however, if  $\mathbf{x}$  is known to be sufficiently *sparse* in some domain, then *exact* recovery of  $\mathbf{x}$  is possible—this is the fundamental tenet of CS theory.

In CS, sparsity is the key aspect that enables recovery of  $\mathbf{x}$  from  $\mathbf{y}$ . Such sparsity can be with respect to some transform  $\Psi$  such that, when the transform is applied to  $\mathbf{x}$ , i.e.,

$$\check{\mathbf{x}} = \Psi \mathbf{x}, \quad (2.2)$$

only  $K$  coefficients in the set of transform coefficients  $\check{\mathbf{x}}$  are nonzero. We say then that  $\mathbf{x}$  is  $K$ -sparse and that  $\Psi$  is the *sparsity basis* for  $\mathbf{x}$ .

We note that real-world signals—particularly the images and the video that we focus on in this dissertation—are rarely, if ever, truly sparse in any basis. Rather, such real-world signals are often only *compressible* in the sense that, if  $\check{x}_n$  are the coefficients of  $\check{\mathbf{x}} = \Psi \mathbf{x}$  sorted in order of decreasing magnitude such that

$$|\check{x}_n| \geq |\check{x}_{n+1}|, \quad (2.3)$$

then

$$|\check{x}_n| < Rn^{-r}, \quad (2.4)$$

where  $r \geq 1$  and  $R < \infty$  (see, e.g., [9, 17, 18, 20]). Define  $\check{\mathbf{x}}_K$  as the set of partial coefficients derived from  $\check{\mathbf{x}}$  by keeping the  $K$  largest coefficients and setting the rest to zero, and  $\mathbf{x}_K = \Psi^{-1}\check{\mathbf{x}}_K$ . Then, if  $\mathbf{x}$  is compressible in the sense of (2.4),  $\mathbf{x}_K$  is close to  $\mathbf{x}$  in the sense of

$$\|\mathbf{x} - \mathbf{x}_K\|_2 \leq C_r R K^{-r}, \quad (2.5)$$

where  $C_r$  is a constant depending only on  $r$  [18]. When the coefficient magnitudes have a power-law decay as in (2.4), CS theory holds that the resulting compressible  $\mathbf{x}$  can be recovered approximately from  $\mathbf{y} = \Phi\mathbf{x}$ . Specifically, the recovered signal,  $\hat{\mathbf{x}}$ , will be close to  $\mathbf{x}_K$  in the sense of

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq C \frac{\|\mathbf{x} - \mathbf{x}_K\|_1}{\sqrt{K}} \quad (2.6)$$

for some well-behaved constant  $C$  (see, e.g., [17, 20]).

In the case of such approximate recovery of compressible signals, the subrate  $S = M/N$  plays the crucial role in determining the quality of the reconstruction—a larger subrate will enable a sparse approximation containing more nonzero coefficients to be recovered, leading to a reconstruction that is closer to the original signal. As a consequence, in experimental results to come later, we will evaluate performance of image- and video-reconstruction strategies by examining an image-distortion measure over a range of subrate values, noting that perfect, distortion-less, recovery is achieved for such compressible signals only when  $S = 1$ .

Finally, it is important to note that CS does not necessarily imply compression in the information-theoretic, rate-distortion sense. That is, the expression “compressed” in the CS name more correctly refers to a process of dimensionality reduction rather than “compres-

sion” in the form of source coding as this term is commonly construed in the information-theory community. Similarly, “compressibility” in the CS domain is strictly in the sense of (2.4), and we emphasize that our use of this term in the remainder of our discussion will be specifically limited to this context.

### 2.1.1 Approaches to CS-Based Signal Acquisition

In the CS framework, the signal-acquisition or sensing device acquires neither  $\mathbf{x}$  nor its coefficients  $\tilde{\mathbf{x}}$  directly. Instead, the sensing device measures linear projections of  $\mathbf{x}$  onto the measurement basis  $\Phi$ , thereby acquiring the signal directly in a reduced dimensionality. This CS measurement process is represented mathematically as (2.1).

The trick behind CS is, of course, to do this dimensionality-reducing signal acquisition in a reversible manner, and this imposes conditions on the measurement and sparsity bases. Specifically, CS theory dictates that it is sufficient that the sparsity basis and the measurement basis be mutually *incoherent* in the sense that the measurement basis  $\Phi$  cannot sparsely represent the columns of the sparsity basis  $\Psi$ . The usual choice for the measurement basis  $\Phi$  is a random matrix, since it can be shown that a random basis will be incoherent with *any* sparsity basis with very high probability. In this sense, a random measurement basis works universally well for any signal regardless of the domain in which its sparsity exists. Random matrices that exhibit such incoherence with any sparsity basis include matrices populated with independent and identically distributed Gaussian or  $\pm 1$

Rademacher<sup>1</sup> random variables [6]. An orthonormalized random matrix (formed by, e.g., orthonormalizing the rows of the above Gaussian matrix) also provides incoherence [22].

We observe that the most straightforward approach to representing the measurement matrix  $\Phi$  is simply an  $M \times N$  dense array of values (e.g., floating point for Gaussian matrices, binary for Rademacher matrices); in fact,  $\Phi$  has been implemented this way in many applications and in much of the CS literature. However, if the dimensionality of  $\mathbf{x}$  is large (as is often the case when multidimensional signals like images and video are considered), the memory required to store  $\Phi$  in both the sensing and reconstruction devices may be impractically large; additionally, a large  $\Phi$  may result in a huge computation burden for the reconstruction process which typically involves a large number of multiplications with  $\Phi$ . However, the use of a structurally random matrix (SRM) (e.g., [35, 51]) can significantly mitigate these issues. In essence, an SRM provides a signal-sensing process (operator  $\Phi$ ) consisting of a random permutation, a simple and computationally efficient transform (such as a block cosine or Hadamard transform), and a random subsampling process. Each of these components can be generated procedurally without requiring the storage of a large, dense matrix; additionally, an SRM can be applied to the input signal with little computation or memory. As a consequence, many CS-reconstruction implementations use SRMs in practice, and it is largely anticipated that large-dimensional CS signal-acquisition devices will as well.

Finally, we argue that CS makes the most sense in situations in which the acquisition of each measurement (i.e., each component of vector  $\mathbf{y}$  in (2.1)) comes with some substantial

---

<sup>1</sup>A Rademacher distribution assigns probability  $\frac{1}{2}$  to the values  $\pm 1$ .

associated cost. For example, in medical applications, each measurement may represent the exposure of the patient to a certain amount of radiation—clearly, signal-acquisition schemes that minimize the total radiation exposure are desirable. On the other hand, in many imaging applications, such as those we consider later, each measurement may represent a discrete sensing element—the fewer the elements used, the lower the overall device cost, particularly for wavelengths outside of the visible spectrum for which a single sensor element may be so costly such that a full array of millions of sensors may be prohibitively expensive.

Under this “measurements-are-costly” paradigm, the potential advantages of CS arise from a signal-sensing process in which  $\mathbf{x}$  is acquired and *simultaneously* reduced in dimension in the form of  $\mathbf{y}$  directly within the hardware of the sensing device. In this case, the matrix-vector multiplication of the measurement process,  $\Phi\mathbf{x}$ , is performed implicitly within the sensor rather than calculated explicitly. In this sense, CS-based signal acquisition offers computation-free dimensionality reduction at the sensor side of the system. In such a sensing device, the full-dimensional signal  $\mathbf{x}$  does not exist at any point within the sensor, having never been sensed or acquired in its full dimensionality.

### 2.1.2 Approaches to CS Reconstruction

Given CS measurements  $\mathbf{y}$  produced via (2.1) with measurement matrix  $\Phi$ , the CS reconstruction problem is to find  $\hat{\mathbf{x}}$  such that  $\mathbf{y} = \Phi\hat{\mathbf{x}}$  (either exactly or approximately) and such that the coefficients  $\tilde{\mathbf{x}} = \Psi\hat{\mathbf{x}}$  are sparse with respect to sparsity basis  $\Psi$ . The



most straightforward formulation of the reconstruction problem searches for the  $\tilde{\mathbf{x}}$  with the smallest  $\ell_0$  norm<sup>2</sup> consistent with the observed  $\mathbf{y}$ ; i.e.,

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_0, \quad \text{such that } \mathbf{y} = \Phi\Psi^{-1}\tilde{\mathbf{x}}, \quad (2.7)$$

where  $\Psi^{-1}$  is the inverse sparsity transform. For the  $\tilde{\mathbf{x}}$  solving (2.7), the final reconstruction is then

$$\hat{\mathbf{x}} = \Psi^{-1}\tilde{\mathbf{x}}. \quad (2.8)$$

However, this  $\ell_0$  optimization is NP-hard (e.g., [21]) and thus computationally infeasible for all but the smallest of problems. As a consequence, there have been a large number of alternative optimizations proposed in recent literature. Perhaps the most prominent of these is basis pursuit (BP) [25] which applies a convex relaxation to the  $\ell_0$  problem resulting in an  $\ell_1$  optimization,

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1, \quad \text{such that } \mathbf{y} = \Phi\Psi^{-1}\tilde{\mathbf{x}}. \quad (2.9)$$

Often, it is assumed that the CS measurements are acquired with some noise; i.e., (2.1) becomes

$$\mathbf{y} = \Phi\mathbf{x} + \mathbf{n}, \quad (2.10)$$

where  $\mathbf{n}$  is some vector of noise. In this case, one can relax the equality constraint in the  $\ell_1$  formulation of (2.9) to yield<sup>3</sup>

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1, \quad \text{such that } \|\mathbf{y} - \Phi\Psi^{-1}\tilde{\mathbf{x}}\|_2^2 \leq \epsilon \quad (2.11)$$

---

<sup>2</sup>The  $\ell_0$  “norm,”  $\|\tilde{\mathbf{x}}\|_0$ , is not truly a norm but is a pseudonorm; it is merely the number of nonzero coefficients in  $\tilde{\mathbf{x}}$ .

<sup>3</sup>The well-known least absolute shrinkage and selection operator (LASSO) [99] has a form similar to (2.11) with the  $\ell_2$  term as the minimization objective and  $\ell_1$  term as the constraint.

for some tolerance  $\epsilon > 0$ . This constrained optimization is closely related to the unconstrained Lagrangian formulation,

$$\min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1 + \lambda \|\mathbf{y} - \Phi\Psi^{-1}\tilde{\mathbf{x}}\|_2^2, \quad (2.12)$$

known as basis-pursuit denoising (BPDN) [25], wherein the Lagrangian multiplier  $\lambda$  balances the  $\ell_1$ -driven sparsity against the  $\ell_2$ -based measure of distortion.

A variety of strategies have been proposed to solve the (2.9), (2.11), and (2.12) optimizations; see [102] for a survey. BP (2.9) and BPDN (2.12) can be solved effectively with linear and second-order-cone programs, respectively; the well-known  $\ell_1$ -MAGIC<sup>4</sup> software package takes this approach, for example. However, such convex-programming methods have computational complexity that is often large for higher-dimensional signals such as images and video. As an alternative, gradient-descent methods solve BPDN (2.12) and tend to be faster in practice than corresponding interior-point solutions. Gradient-based algorithms include iterative splitting and thresholding (IST) [28], sparse reconstruction via separable approximation (SpaRSA) [110], and gradient projection for sparse reconstruction (GPSR) [46].

A number of greedy algorithms have also been proposed for the CS reconstruction problem. These include matching pursuits [80], orthogonal matching pursuits (OMP) [101], and compressive sampling matching pursuit (CoSaMP) [86]. In practice, such greedy algorithms significantly reduce computational complexity as compared to convex-programming approaches, albeit typically at the cost of lower reconstruction quality [102].

---

<sup>4</sup><http://www.l1-magic.org>

Iterative thresholding (e.g., [13, 14, 28, 57]) is a body of algorithms which serves as an alternative to the greedy pursuits class of CS reconstruction. Iterative-thresholding algorithms form  $\tilde{\mathbf{x}}$  by successively projecting and thresholding; for example, the reconstruction in [57] starts from some initial approximation  $\tilde{\mathbf{x}}^{(0)}$  and forms the approximation at iteration  $i + 1$  as

$$\tilde{\mathbf{x}}^{(i)} = \tilde{\mathbf{x}}^{(i)} + \frac{1}{\gamma} \Psi \Phi^T (\mathbf{y} - \Phi \Psi^{-1} \tilde{\mathbf{x}}^{(i)}), \quad (2.13)$$

$$\tilde{\mathbf{x}}^{(i+1)} = \begin{cases} \tilde{\mathbf{x}}^{(i)}, & |\tilde{\mathbf{x}}^{(i)}| \geq \tau^{(i)}, \\ 0 & \text{else.} \end{cases} \quad (2.14)$$

Here,  $\gamma$  is a scaling factor, while  $\tau^{(i)}$  is a threshold set appropriately at each iteration. For  $\gamma$ , [57] proposes the largest eigenvalue of  $\Phi^T \Phi$ ; in this case, if  $\Phi$  is an orthonormal matrix,  $\Phi^T \Phi = \mathbf{I}$ , and  $\gamma = 1$ . It is straightforward to see that this procedure is a specific instance of a projected Landweber (PL) algorithm [11, 74]; similar approaches for CS reconstruction include [13, 14]. Like the greedy algorithms of the pursuits class, PL-based CS reconstruction also provides reduced computational complexity as compared to convex-programming-based reconstruction. Additionally, the PL formulation offers the possibility of easily incorporating additional optimization criteria. For example, later in this dissertation, we will overview an image-reconstruction technique that incorporates Wiener filtering into the PL iteration to search for a CS reconstruction simultaneously achieving sparsity and smoothness.

The discussion up to now is intended largely as a brief overview of CS theory and the generic algorithms used for CS reconstruction. Next, we turn our attention to the specific

issue of the CS of 2D still images which is the issue at the heart of the remainder of this discussion.

## **2.2 CS Acquisition of Still Images**

There has been significant interest in CS in systems that acquire and process 2D still images. Widespread use of digital cameras has led to increasing demands for higher spatial resolution, lower power consumption, and lower overall device costs. In many existing digital cameras, images are acquired with several million sensor elements. CS offers a compelling alternative to this traditional image-acquisition paradigm—instead of sampling in high resolution, CS offers the possibility of directly acquiring the image in a reduced dimensionality. With this dimensionality reduction taking place implicitly within the hardware of the sensing device, it is speculated that CS may eventually yield camera architectures that are significantly cheaper and that consume less power, both as a result of using a number of discrete sensing elements that is greatly reduced as compared to the full sensor array. Such cameras may then be able to accommodate spectral wavelengths (e.g., infrared) for which a single sensing element is so costly that a multi-mega-sensor array is prohibitive. Cameras might also be tailored to wavelengths impossible to implement in conventional CCD and CMOS imagers [98].

### **2.2.1 The Single-Pixel Camera**

Perhaps the most well-known architecture for the CS acquisition of still images is the so-called “single-pixel camera” [41, 98, 106, 107] developed at Rice University. This camera architecture is illustrated in Figure 2.1. In short, the single-pixel camera uses a digital

micromirror device (DMD) to optically perform inner products in the CS measurement process.

Specifically, the DMD—normally an output device used in digital-light-processing (DLP) applications—forms a pixel array of electrostatically actuated mirrors that can each be oriented in one of two states ( $+12^\circ$  or  $-12^\circ$  from horizontal). When coupled with an analog photosensor, the DMD is used as an image sensor such that the image in question is focused onto the DMD and partially reflected toward the photosensor. By orienting the photosensor such that one of the DMD mirror states reflect light toward the photosensor while the other mirror state reflects away, the DMD, in effect, forms an inner product between the image being acquired and the binary pattern present on the DMD array. That is, the photosensor senses the analog sum of all the light reflected to it from the DMD, thereby outputting a continuous-valued analog measurement for each pattern on the DMD.

The resulting analog voltage reading of the photosensor provides a single (scalar) CS measurement value,  $y_m \in \mathbb{R}$ . By repeating this process  $M$  times with multiple pseudo-random DMD patterns, a vector,  $\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_M \end{bmatrix} \in \mathbb{R}^M$ , of  $M$  CS measurements is acquired. It is straightforward for the binary DMD patterns to effectuate a pseudorandom Rademacher  $\pm 1$  measurement basis by taking an additional measurement with all the mirrors reflecting toward the sensor and then subtracting this mean measurement,  $\bar{y}$ , from the other measurements:

$$2\mathbf{y} - \bar{y} \cdot \mathbf{1} \rightarrow \mathbf{y}, \quad (2.15)$$

where  $\mathbf{1}$  is an  $M \times 1$  vector of all ones [41, 98]. Alternatively, by appropriately duty-cycling the mirrors, a Gaussian measurement matrix can be obtained [41, 98]. As mentioned previ-

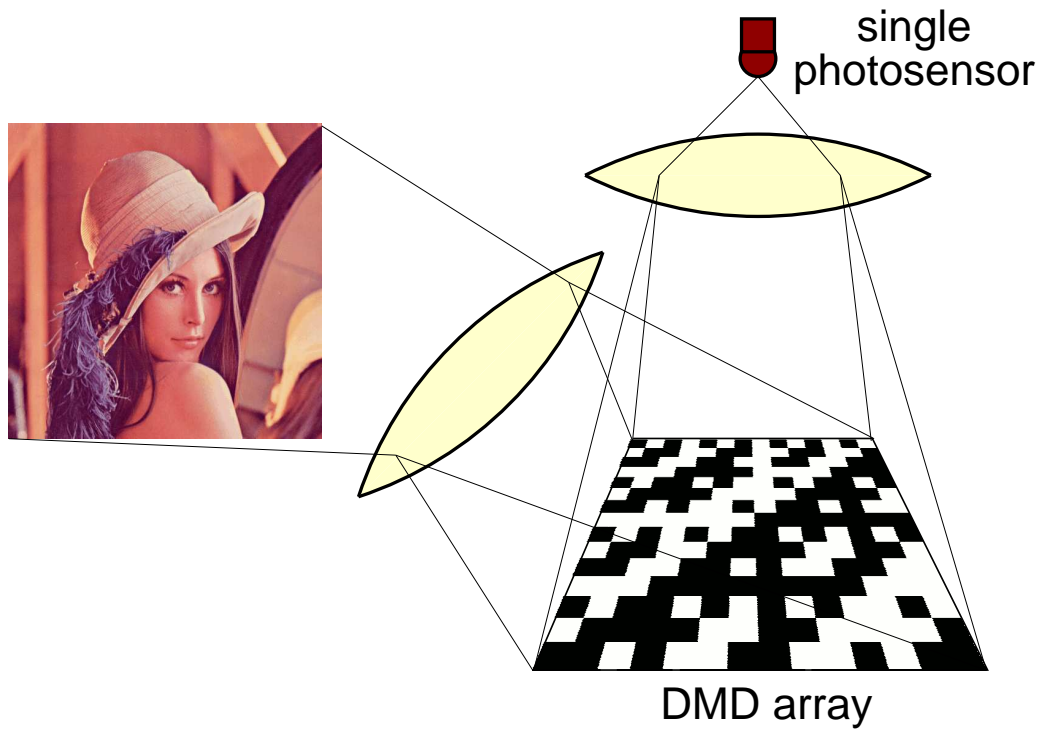


Figure 2.1

The single-pixel camera [41, 98, 106, 107] for the CS acquisition of a still image (figure is from [49]).

ously, such Rademacher and Gaussian measurement bases are incoherent with any sparsity basis with high probability.

Sampling in the single-pixel camera may also be driven by a SRM employing a block Hadamard transform as suggested in [51]. In this latter case, the DMD mirror patterns do not need to be explicitly stored in the sensing device; rather, they are generated procedurally “on the fly” as needed by the measurement process, thereby significantly reducing memory requirements on-board the sensing device as compared to dense Rademacher or Gaussian matrices, particularly when the image size is large.

The CS-based single pixel camera has the potential to provide substantial benefits for hardware-constrained systems but does come at a cost when compared with traditional systems: acquisition time. A CS device requires less memory to store sampled signals and less power to transmit them, as well no computation required to calculate the dimensionality reduction (which is, in effect, accomplished optically), but the total time to acquire a given signal is increased by a factor of  $M$ —more, if duty cycling for a Gaussian matrix is needed (note, however, that such duty cycling is not required for a Hadamard-based SRM as in [51]). For the imaging of static scenes, this trade off is nominal, and the ability to use a single sensor instead of an entire array of sensors might outweigh the cost of the additional exposure time required for each measurement. For dynamic scenes, CS sampling is more complicated (see [49]).

We note also that the single-pixel-camera architecture is perhaps best suited to applications at wavelengths outside of the visible spectrum. Indeed, current CCD and CMOS imagers already in widespread use in digital cameras provide a full array consisting of many

millions of sensor elements in the visible spectrum—since such CCD/CMOS devices are extremely cheap, fast, low-powered, and high-quality, it may be difficult to justify the use of a single-pixel camera in visible-light applications. In contrast, however, sensors for extra-visible wavelengths typically entail significantly higher cost—for example, current infrared cameras are at least a thousand-fold more expensive than visible-light cameras at a comparable spatial resolution. On the other hand, the single-pixel camera, which requires only a single sensing element, may have the potential to significantly reduce the hardware costs of image acquisition in infrared and other extra-visible applications.

### **2.2.2 Other Image-Acquisition Architectures**

In addition to the single-pixel camera, several other architectures for the CS acquisition of still imagery have been proposed; an overview of some of these devices is presented in [42]. These alternative architectures include devices which couple a sensor array with explicit calculation of a random projection in analog hardware. For example, [66] implements random convolution (convolution with a random filter followed by downsampling [89]) for CS image acquisition; on the other hand, [88] performs analog vector-matrix multiplication to effectuate CS measurements. Both [66] and [88] have the drawback that they require implementation of a full array of sensing elements (i.e., equal in number to the pixels in the image) and thus do not reduce the cost associated with the number of such sensing elements. Rather, the assumption is that the analog calculation of the dimensionality reduction is cheaper—in terms of device implementation or power consumption—than other means for accomplishing the same (e.g., digital calculation of an image transform).



Such an advantage has not yet been clearly demonstrated for such devices to the best of our knowledge, however.

Alternatively, [81] proposes CS image acquisition via a coded aperture and a subsampled focal plane. Like the single-pixel camera, a reduced number of sensing elements is used as compared to the full sensor array, but still significantly more than just a single sensor as in the single-pixel camera. Like the single-pixel camera, coded-aperture image acquisition is likely to be most suited to applications residing outside the visible spectrum since current full-array CCD and CMOS imagers already provide low-cost image acquisition in the visible spectrum. Yet extra-visible wavelengths may require significantly more expensive sensor elements and therefore may benefit from the reduced number of sensor elements required by coded-aperture sensors.

Throughout the remainder of this work, we will assume that the single-pixel-camera architecture of Figure 2.1 is employed for CS acquisition of a still image, either through use of a duty-cycled Gaussian measurement matrix or an SRM-driven sampling process. The primary concern is then how one reconstructs the image from the resulting CS measurements; we explore this issue next.

### 2.3 Straightforward Reconstruction for Images

A straightforward implementation of CS on 2D images recasts the 2D array-based problem as a 1D vector-based problem, typically using the following procedure:

1. “Rasterize”  $N \times N$  image  $\mathbf{X}$  into an  $N^2$ -dimensional vector  $\mathbf{x}$ :

$$\mathbf{x} = \text{Raster}(\mathbf{X}). \quad (2.16)$$

Here,  $\text{Raster}(\cdot)$  is a suitable rasterization operator—for example, the concatenation of the  $N$  rows of the image together, followed by a transpose, to produce a  $N^2 \times 1$  vector.

2. Apply  $M \times N^2$  measurement matrix  $\Phi$ :

$$\mathbf{y} = \Phi \mathbf{x}. \quad (2.17)$$

3. Apply some CS reconstruction algorithm. This vector-based reconstruction will employ  $\Phi$ , the measurement matrix, as well as  $\Psi$ , the sparsity transform. In this context,  $\Psi$  is an  $N^2 \times N^2$  matrix consisting of the  $N^2$  basis images of some 2D image transform (discrete cosine transform (DCT), discrete wavelet transform (DWT), etc.), arranged such that each basis image has been rasterized into a  $N^2 \times 1$  vector and placed columnwise into  $\Psi$ . The reconstruction yields  $\hat{\mathbf{x}} = \Psi \tilde{\mathbf{x}}$ :

$$\hat{\mathbf{x}} = \text{CS\_Reconstruction}(\mathbf{y}, \Phi, \Psi). \quad (2.18)$$

4. Reassemble the image from the vectorized reconstruction:

$$\hat{\mathbf{X}} = \text{Unraster}(\hat{\mathbf{x}}). \quad (2.19)$$

This rasterized paradigm was used, for example, for reconstruction with the single-pixel-camera system in [98, 106, 107].

We note that the description here of the sparsity basis  $\Psi$  as an  $N^2 \times N^2$  matrix is largely conceptual. While it would certainly be possible to represent any 2D image transform such as a DCT or DWT in this fashion, in practical implementation of CS reconstruction, one would almost certainly use some fast computation of both the forward transform  $\Psi \tilde{\mathbf{x}}$  and the inverse transform  $\Psi^{-1} \mathbf{x}$  within the  $\text{CS\_Reconstruction}(\cdot)$  process. For example, there are a number of fast algorithms for both the DCT (e.g., [45, 78]) and DWT (e.g., [30, 97]). Such algorithms typically offer a number of advantages for the calculation of an image transform as compared to matrix-vector multiplication, which, although conceptually simple, can entail significantly more computation and memory than the fast algorithms.

A primary concern with (2.18) is that, when a generic CS reconstruction is employed, it is somewhat “blind” to the fact that the data being processed represents an image beyond the fact that the sparsity transform  $\Psi$  is a 2D image transform. That is, generic CS reconstruction algorithms treat the problem as any other, simply searching for a sparse solution that is consistent with the observed random measurements. Unfortunately, imposing sparsity alone does not necessarily produce the most visually pleasing reconstructed images. For example, [17] found that the sparse solutions produced by a straightforward CS reconstruction often incurred visually unpleasant, high-frequency oscillations. In essence, the issue stems from the fact that generic CS reconstruction ignores attributes known to be widely possessed by images, such as smoothness. We will examine strategies for incorporating such image attributes into CS reconstruction below.

Another concern is that, from the perspective of practical implementation, the fact that the size of  $\Phi$  is  $O(N^4)$  entails that the memory required to store this matrix grows very fast as the number of pixels in the image in question increase. This leads to a huge memory required to store the measurement operator when  $\Phi$  is implemented as a dense matrix (e.g., Gaussian or Rademacher) within the CS sensing process. Additionally, a large  $\Phi$  yields a huge memory and computational burden within the CS reconstruction process as well, since CS reconstruction involves numerous matrix computations with  $\Phi$ . As a consequence of these difficulties in both memory and computation, many instances of image CS in prior literature consider images no larger than  $256 \times 256$  pixels, and some focus on even smaller sizes. As mentioned previously, SRMs (e.g., [35, 51]) can effectively reduce these computation and memory burdens by using a procedurally-generated measurement opera-

tor in both the sensing and reconstruction processes. However, an alternative to SRMs is to impose some form of sparse structure (e.g., block diagonality) onto  $\Phi$ , a possibility that we explore in detail below.

We now consider several strategies that attempt to ameliorate the difficulties identified above. Specifically, we first consider a reconstruction based on total variation that capitalizes on the fact that the underlying data to be reconstructed represents a natural 2D image. A primary focus of this effort is the promotion of smoothness of the end reconstruction. We then consider CS based on image blocks that effectively focuses on both smoothness of reconstruction as well as fast and memory-efficient implementation of the measurement operator.

## 2.4 Total-Variation Reconstruction

In [17], BPDN as given in (2.12) is recast in order to promote smoothness in the reconstructed images and to suppress the high-frequency artifacts encountered with the straightforward solution to (2.12). Specifically, it is proposed that, instead of seeking sparsity in the domain of some image transform, (2.12) is reformulated as

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{\text{TV}} + \lambda \|\mathbf{y} - \Phi \mathbf{x}\|_2, \quad (2.20)$$

where  $\mathbf{x}$  is the 1D rasterization of image  $\mathbf{X}$ . Here, the total variation (TV) of the image is

$$\|\mathbf{X}\|_{\text{TV}} = \sum_{i,j} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}, \quad (2.21)$$

where  $x_{i,j}$  is the pixel in location  $(i, j)$  in the image  $\mathbf{X}$ . In [17], (2.20) is solved using a second-order-cone program that accommodates the TV-based norm.

In essence, the TV approach to CS reconstruction replaces sparsity in the domain of an image transform with sparsity in the domain of a discretized gradient, implicitly promoting smoothness in the resulting reconstruction. As a result, TV-based reconstruction ameliorates the problem associated with generic CS recovery of image data in that the reconstruction is no longer blind to the fact that the underlying signal is an image. Although the use of SRMs [35, 51] can greatly improve matters, TV reconstruction still tends to be very computationally complex as compared to other reconstruction algorithms. As a consequence, TV-based CS reconstruction of an image tends to be rather slow, perhaps even to the point of being infeasible if the image size is large.

## 2.5 CS with Blocks in the Spatial Domain

As an alternative to SRMs for alleviating the huge computation and memory burdens associated with a dense measurement matrix  $\Phi$  within both the sensing and reconstruction processes, one can adopt a philosophy long used in image-processing fields when an image is too large to be feasibly processed in its entirety—specifically, break the image into smaller blocks and process the blocks independently. Such an approach was proposed in [50] for block-based compressed sensing (BCS) for 2D images.

In BCS, an image is divided into  $B \times B$  non-overlapping blocks and acquired using an appropriately-sized measurement matrix. That is, suppose that  $\mathbf{x}_j$  is a vector representing, in raster-scan fashion, block  $j$  of input image  $\mathbf{X}$ . The corresponding  $\mathbf{y}_j$  is then

$$\mathbf{y}_j = \Phi_B \mathbf{x}_j, \quad (2.22)$$

where  $\Phi_B$  is an  $M_B \times B^2$  measurement matrix such that the target substrate for the image as a whole is  $S = M_B/B^2$ . It is straightforward to see that (2.22) applied block-by-block to an image is equivalent to a whole-image measurement matrix  $\Phi$  in (2.17) with a constrained structure; specifically,  $\Phi$  is block diagonal,

$$\Phi = \begin{bmatrix} \Phi_B & 0 & \cdots & 0 \\ 0 & \Phi_B & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \Phi_B \end{bmatrix}. \quad (2.23)$$

We note that the single-pixel camera discussed above can easily accommodate BCS acquisition by simply driving the DMD array with this block-diagonal  $\Phi$  instead of a dense  $\Phi$  as was done originally in [41, 98, 106, 107].

There are several approaches that one can take to reconstruct an image that has been acquired using BCS. The most straightforward situation is when the sparsity transform  $\Psi$  is also a block-based operator of the same  $B \times B$  size, i.e.,

$$\Psi = \begin{bmatrix} \Psi_B & 0 & \cdots & 0 \\ 0 & \Psi_B & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \Psi_B \end{bmatrix}, \quad (2.24)$$

where  $\Psi_B$  is a  $B^2 \times B^2$  matrix for the block-based transform. In this case, (2.18) can simply be applied for each block independently; i.e., for block  $j$ ,

$$\hat{\mathbf{x}}_j = \text{CS\_Reconstruction}(\mathbf{y}_j, \Phi_B, \Psi_B). \quad (2.25)$$

We refer to this independent, block-by-block reconstruction as a “block-independent” reconstruction. In general, block-independent reconstruction will produce severe blocking artifacts and is thus not usually a reasonable solution.

Better results may arise when the block measurement matrix  $\Phi_B$  is placed into the block-diagonal  $\Phi$  as in (2.23) with  $\Phi$  then being used in the whole-image reconstruction of (2.18). We refer to this as “block-diagonal” reconstruction. Block-diagonal reconstruction permits the sparsity achieved in the reconstructed signal to be adjusted across all blocks simultaneously (in the block-independent solution, sparsity is imposed on each block independently). In this case, the sparsity transform  $\Psi$  can take the form of a full-image transform (such as a DWT) rather than being constrained to have a block-diagonal structure as is the case for block-independent reconstruction. Such a full-image transform will serve to ameliorate the blocking artifacts that naturally arise in BCS due to the block-diagonal structure of  $\Phi$ . More importantly, since block-based transforms do not decorrelate across image blocks, the compressibility of the image in the sense of (2.4) is likely to be much greater in the domain of a full-image transform such as a DWT, achieving a much higher-fidelity reconstruction than the block-independent approach.

An alternative is to eliminate the sparsity transform  $\Psi$  altogether and opt instead to impose a sparse gradient, i.e., by adopting a TV-based reconstruction. In such a TV-based solution, we apply (2.20) using the block-diagonal  $\Phi$  of (2.23). We refer to this combination of BCS and TV-based reconstruction as “BCS-TV.”

In experimental results to follow, we will evaluate the performance of each of these strategies and give examples of the visual quality of reconstructed images. First, however,

we overview an alternative to BCS-TV that combines BCS acquisition, fast iterative reconstruction, and explicit smoothing in the form of Wiener filtering with the goal of producing high-quality, visual-pleasing reconstructions like BCS-TV without its heavy computational burden.

### 2.5.1 The BCS-SPL Algorithm

In [50], BCS was proposed wherein the acquisition of an image is driven by random matrices applied on a block-by-block basis, while the reconstruction is a variant of the PL reconstruction of (2.13)–(2.14) that incorporates a smoothing operation intended to reduce blocking artifacts. Since it combines BCS with a smoothed projected Landweber (SPL) reconstruction, in [82], the overall technique was called BCS-SPL.

In BCS-SPL, Wiener filtering is incorporated into the basic PL framework in order to remove blocking artifacts. In essence, this operation imposes smoothness in addition to the sparsity inherent to PL. Specifically, a Wiener-filtering step is interleaved with the PL projection of (2.13)–(2.14). The specific implementation we use here was initially described in [82] and is presented in Figure 2.2, where  $\text{Wiener}(\cdot)$  is pixel-wise adaptive Wiener filtering using a neighborhood of  $3 \times 3$ ,  $\text{Threshold}(\cdot)$  is a thresholding process, and  $\Phi_B$  is assumed to be a random orthonormal matrix such that  $\gamma$  in (2.13) is unity.

In [82], to set a proper  $\tau$  for hard thresholding in  $\text{BCS-SPL}(\cdot)$ , we employ the so-called interference heuristic [79] which effectively assumes Gaussian marginal statistics for the insignificant transform coefficients (i.e., those coefficients that are zero in a sparse approx-



```

function  $\mathbf{x} = \text{BCS-SPL}(\mathbf{y}, \Phi_B, \Psi)$ 
  for each block  $j$ 
     $\mathbf{x}_j^{(0)} = \Phi_B^T \mathbf{y}_j$ 
   $i = 0$ 
  do
     $\hat{\mathbf{x}}^{(i)} = \text{Wiener}(\mathbf{x}^{(i)})$ 
    for each block  $j$ 
       $\hat{\mathbf{x}}_j^{(i)} = \hat{\mathbf{x}}_j^{(i)} + \Phi_B^T (\mathbf{y}_j - \Phi_B \hat{\mathbf{x}}_j^{(i)})$ 
     $\check{\mathbf{x}}^{(i)} = \Psi \hat{\mathbf{x}}^{(i)}$ 
     $\check{\mathbf{x}}^{(i)} = \text{Threshold}(\check{\mathbf{x}}^{(i)})$ 
     $\bar{\mathbf{x}}^{(i)} = \Psi^{-1} \check{\mathbf{x}}^{(i)}$ 
    for each block  $j$ 
       $\mathbf{x}_j^{(i+1)} = \bar{\mathbf{x}}_j^{(i)} + \Phi_B^T (\mathbf{y}_j - \Phi_B \bar{\mathbf{x}}_j^{(i)})$ 
     $D^{(i+1)} = \|\mathbf{x}^{(i+1)} - \hat{\mathbf{x}}^{(i)}\|_2$ 
     $i = i + 1$ 
  until  $|D^{(i)} - D^{(i-1)}| < 10^{-4}$ 
   $\mathbf{x} = \mathbf{x}^{(i)}$ 

```

Figure 2.2

BCS-SPL reconstruction of a 2D image.

imation to the original image). Then, the universal threshold of [36] is used. Specifically, in (2.14),

$$\tau^{(i)} = \lambda \sigma^{(i)} \sqrt{2 \log K}, \quad (2.26)$$

where  $\lambda$  is a constant control factor to manage convergence, and  $K$  is the number of the transform coefficients. As in [36],  $\sigma^{(i)}$  is estimated using a robust median estimator,

$$\sigma^{(i)} = \frac{\text{median}(|\check{\mathbf{x}}^{(i)}|)}{0.6745}. \quad (2.27)$$

As a final note, we observe that while (2.26) and (2.27) form a convenient heuristic for setting the threshold in the BCS-SPL algorithm, [38] points to theoretical shortcomings with the underlying Gaussian assumption in general iterative-thresholding algorithms for CS.

### 2.5.2 BCS-SPL with Directional Transforms

The BCS-SPL framework of [50, 82] is quite flexible thanks to its rather simple implementation. It is straightforward to incorporate sophisticated transforms and thresholding, as well as additional constraints into the process. As a specific contribution of this dissertation, we incorporate highly directional transforms with statistically estimated thresholding into the BCS-SPL framework, work that we first proposed in [82] and then further expanded upon in [49].

Although DWTs are widely used in traditional source coding of images (e.g., the JPEG-2000 standard [63]), DWTs in their traditional critically-sampled form are known to be somewhat deficient in several characteristics, lacking such properties as shift invariance and significant directional selectivity. As a result, there have been several recent proposals

made for transforms that feature a much higher degree of directional representation than is obtained with traditional DWTs. Two prominent families of such directional transforms are contourlets and complex-valued DWTs. The contourlet transform (CT) [32] preserves interesting features of the traditional DWT, namely multiresolution and local characteristics of the signal, and, at the expense of spatial redundancy, it better represents the directional features of the image. The CT couples a Laplacian-pyramid decomposition with directional filterbanks, inheriting the redundancy of the Laplacian pyramid (i.e.,  $4/3$ ). Alternatively, complex-valued wavelet transforms have been proposed to improve upon DWT deficiencies, with the dual-tree discrete wavelet transform (DDWT) [73] becoming a preferred approach due to the ease of its implementation. In the DDWT, real-valued wavelet filters produce the real and imaginary parts of the transform in parallel decomposition trees. The DDWT yields a decomposition with a much higher degree of directionality than that possessed by traditional DWTs; however, since both trees of the DDWT are themselves orthonormal or biorthogonal decompositions, the DDWT taken as a whole is a redundant tight frame. Albeit redundant, both the CT and DDWT have been effectively used in the source coding of images (e.g., [15, 23, 43, 44, 47, 100]); their use in BCS-SPL reconstruction was first proposed in [82].

Hard thresholding inherently assumes independence between coefficients. However, bivariate shrinkage [26] is better suited to directional transforms in that it exploits statistical dependency between transform coefficients and their respective parent coefficients, yielding performance superior to that of hard thresholding. In [26], a non-Gaussian bivariate distribution was proposed for the current coefficient and its lower-resolution parent

coefficient based on an empirical joint histogram of DWT coefficients. However, it is straightforward to apply this process to any transform having a multiple-level decomposition, such as the directional transforms we consider here. Specifically, we propose that, given a specific transform coefficient  $\xi$  and its parent coefficient  $\xi_p$  in the next coarser scale, the  $\text{Threshold}(\cdot)$  operator in SPL be the maximum *a posteriori* estimator of  $\xi$ ,

$$\text{Threshold}(\xi, \lambda) = \frac{\left(\sqrt{\xi^2 + \xi_p^2} - \lambda \frac{\sqrt{3}\sigma^{(i)}}{\sigma_\xi}\right)_+}{\sqrt{\xi^2 + \xi_p^2}} \cdot \xi, \quad (2.28)$$

where  $(g)_+ = 0$  for  $g < 0$ ,  $(g)_+ = g$  else;  $\sigma^{(i)}$  is the median estimator of (2.27); and, again,  $\lambda$  is a convergence-control factor. Here,  $\sigma_\xi^2$  is the marginal variance of coefficient  $\xi$  estimated in a local  $3 \times 3$  neighborhood surrounding  $\xi$  as in [26]. We refer the reader to [26] for full details on the method of bivariate shrinkage and its underlying statistical models.

### 2.5.3 Experimental Observations on BCS-Based Reconstruction

As another primary contribution of this dissertation, we now present a battery of experimental evaluations that gauge the performance of the block-based still-image reconstruction techniques just previously discussed. Throughout, we employ several popular grayscale images<sup>5</sup> of size  $512 \times 512$ , and we use BCS on blocks of size  $32 \times 32$ . In all cases, the images are subjected to a BCS measurement process with  $\Phi_B$  in (2.22) being an orthonormalized dense Gaussian matrix. We measure reconstruction performance in terms of a peak signal-to-noise ratio (PSNR) between the reconstructed and original images. We note that the results presented here were previously reported in [49].

---

<sup>5</sup>The images are available in the BCS-SPL package, <http://www.ece.mstate.edu/~fowler/BCSSPL/>

We compare several reconstruction techniques as discussed above. Figures 2.3–2.4 depict typical reconstruction results for the “Lenna” image using a subrate  $S = 0.2$ ; these figures show only a detailed portion from the center of the image. Figure 2.3(a) gives the original image, while Figures 2.3(b)–(d) present several straightforward reconstructions, all based on a generic GPSR CS reconstruction algorithm [46]. The first, in Figure 2.3(b), comes from a block-independent reconstruction using a block-based DCT as the sparsity transform; in this “block-independent BCS-GPSR-DCT” reconstruction, each  $32 \times 32$  block is reconstructed independently via (2.25) with  $\Psi_B$  being the block-based DCT. Figure 2.3(c) presents a similar reconstruction, only this time, the reconstruction is block-diagonal. That is, “block-diagonal BCS-GPSR-DCT” instead performs whole-image reconstruction via (2.18) using the corresponding block-diagonal  $\Phi$  and  $\Psi$ , the latter again being based on the  $32 \times 32$  block DCT. Severe blocking artifacts are readily apparent in Figures 2.3(b) and (c). This blocking is ameliorated significantly in Figure 2.3(d) in which a whole-image DWT is used in the reconstruction via (2.18); i.e., in BCS-GPSR-DWT,  $\Phi$  still possesses a block-diagonal structure, but  $\Psi$  does not. We note that PSNR figures for the three straightforward GPSR-based reconstructions are fairly similar, being around 26 dB for “Lenna.”

Significantly improved PSNR performance, as well as visual quality, results when smoothing is included in the reconstruction process. To wit, Figures 2.4(a)–(c) present example reconstructions for a TV-based reconstruction as well as two BCS-SPL-based reconstructions. Specifically, Figure 2.4(a) illustrates the performance of BCS-TV wherein the TV reconstruction of (2.20) is applied to BCS-acquired imagery, promoting a sparse

image gradient and thereby an implicitly smooth reconstruction. We note that we use  $\ell_1$ -MAGIC<sup>6</sup> in the BCS-TV implementation. On the other hand, Figures 2.4(b) and (c) show the reconstruction due to BCS-SPL which features explicit smoothing in the form of a Wiener filter incorporated into the iterative PL reconstruction process. Shown are results for both a  $32 \times 32$  block DCT (BCS-SPL-DCT, Figure 2.4(b)) and a full-frame DWT (BCS-SPL-DWT, Figure 2.4(c)). We use the BCS-SPL implementation available from the BCS-SPL website<sup>7</sup>. The results of the smoothing-based reconstructions of Figures 2.4(a)–(c) are similar, with PSNRs around 30.5 dB, some 4 dB higher than the straightforward GPSR-based reconstructions of Figures 2.3(b)–(d).

Figures 2.4(d) and (e) evaluate directional transforms for CS reconstruction, deploying both the CT and DDWT within the BCS-SPL framework. We refer to the resulting implementations as BCS-SPL-CT and BCS-SPL-DDWT, respectively. The results for the directional transforms are similar, with PSNRs around 31 dB being about 0.5 dB higher than the reconstructions of Figure 2.4.

We note that, in these results, we use bivariate shrinkage (2.28) with  $\lambda = 10, 25$ , and 25, respectively, for BCS-SPL-CT, BCS-SPL-DDWT, and BCS-SPL-DWT. Lacking parent-child relations, BCS-SPL-DCT uses hard thresholding (2.26) with  $\lambda = 6$ .

Table 2.1 compares PSNR for several  $512 \times 512$  images at several subrates. We note that, since the quality of reconstruction can vary due to the randomness of the measurement matrix  $\Phi$ , all PSNR figures are averaged over 5 independent trials. The results indicate that BCS-SPL with the directional transforms achieves the best performance at low measure-

---

<sup>6</sup><http://www.l1-magic.org>

<sup>7</sup><http://www.ece.mstate.edu/~fowler/BCSSPL/>



(a) Original image



(b) Block-independent  
BCS-GPSR-DCT, 25.7 dB



(c) Block-diagonal  
BCS-GPSR-DCT, 25.9,dB



(d) BCS-GPSR-DWT, 26.0 dB

Figure 2.3

Reconstructions of the  $512 \times 512$  “Lenna” image (shown in detail) for a subrate of  $S = 0.2$ . BCS measurement uses  $32 \times 32$  blocks.



Figure 2.4

Reconstructions of the  $512 \times 512$  “Lenna” image (shown in detail) for a subrate of  $S = 0.2$ . BCS measurement uses  $32 \times 32$  blocks.



Table 2.1

BCS-based reconstruction performance as measured in PSNR in dB; figures averaged over 5 independent trials; images are  $512 \times 512$ .

<i>Algorithm</i>	<i>Subrate</i>				
	0.1	0.2	0.3	0.4	0.5
Lenna					
BCS-SPL-DDWT	<b>28.3</b>	<b>31.4</b>	<b>33.5</b>	<b>35.2</b>	<b>36.8</b>
BCS-SPL-CT	28.2	31.0	33.0	34.7	36.3
BCS-SPL-DWT	27.8	30.9	32.9	34.6	36.2
BCS-SPL-DCT	27.7	30.5	32.5	34.2	35.8
BCS-TV	27.9	30.6	32.6	34.3	35.9
BCS-GPSR-DWT	22.7	26.0	28.1	29.9	31.3
Block-diagonal BCS-GPSR-DCT	22.5	25.9	28.1	29.8	31.2
Block-independent BCS-GPSR-DCT	22.6	25.8	28.0	29.8	31.3
Barbara					
BCS-SPL-DDWT	<b>22.9</b>	24.3	<b>25.9</b>	<b>27.5</b>	29.1
BCS-SPL-CT	22.8	24.3	<b>25.9</b>	<b>27.5</b>	<b>29.4</b>
BCS-SPL-DWT	22.6	23.9	25.2	26.6	28.1
BCS-SPL-DCT	22.8	<b>24.4</b>	<b>25.9</b>	27.4	29.1
BCS-TV	22.5	23.6	24.6	25.6	26.7
BCS-GPSR-DWT	20.1	22.2	23.6	25.0	26.4
Block-diagonal BCS-GPSR-DCT	20.3	23.4	25.4	27.3	28.8
Block-independent BCS-GPSR-DCT	20.4	23.3	25.5	27.3	28.9
Mandrill					
BCS-SPL-DDWT	<b>22.9</b>	24.9	26.7	28.4	30.3
BCS-SPL-CT	<b>22.9</b>	<b>25.0</b>	<b>27.0</b>	<b>28.9</b>	<b>30.9</b>
BCS-SPL-DWT	22.5	24.3	26.0	27.7	29.4
BCS-SPL-DCT	22.3	24.2	25.9	27.8	29.7
BCS-TV	22.3	24.3	26.1	27.8	29.5
BCS-GPSR-DWT	18.6	19.8	20.9	21.9	23.1
Block-diagonal BCS-GPSR-DCT	18.5	19.8	20.9	22.1	23.2
Block-independent BCS-GPSR-DCT	18.5	19.8	20.9	22.1	23.3
Goldhill					
BCS-SPL-DDWT	<b>27.0</b>	28.9	30.5	31.8	33.1
BCS-SPL-CT	26.9	<b>29.0</b>	30.5	31.9	33.3
BCS-SPL-DWT	26.7	28.7	30.1	31.5	32.9
BCS-SPL-DCT	26.1	28.3	29.6	31.0	32.6
BCS-TV	26.5	28.9	<b>30.6</b>	<b>32.1</b>	<b>33.6</b>
BCS-GPSR-DWT	23.0	25.4	27.0	28.4	29.6
Block-diagonal BCS-GPSR-DCT	23.1	25.6	27.3	28.6	29.8
Block-independent BCS-GPSR-DCT	22.9	25.5	27.2	28.7	29.9

ment rates. At higher measurement rates, performance is more varied—BCS-TV is more competitive; however, the directional BCS-SPL techniques usually produce PSNR close to that of the TV-based algorithm.

The experimental results presented here demonstrate that imposing smoothness to the reconstruction—either implicitly via the gradient-based TV approach, or explicitly in the form BCS-SPL’s Wiener filtering—can improve PSNR and visual-quality performance significantly. BCS-SPL is advantageous in that its simple formulation based on Landweber iterations permits it to easily accommodate various sophistications such as directional transforms and bivariate shrinkage.

## **2.6 CS with Blocks in the Wavelet-Domain**

As discussed above, techniques such as TV and BCS-SPL exploit smoothness that is anticipated to be present in the original image signal. However, other image characteristics can also be used to tailor CS reconstruction specifically to image data. For example, a number of recent CS strategies (e.g., [56, 58, 59, 72, 90]) are deployed assuming that the image is both acquired and reconstructed in the domain of a DWT. Such wavelet-domain CS permits known statistical models (e.g., [58, 59, 72]) for wavelet coefficients to be exploited in reconstruction. Additionally, the degree of CS subsampling can be adapted to the wavelet decomposition—often, the baseband is retained in full with no subsampling (e.g., [56, 90]), while the degree of subsampling is increased for successively higher-resolution decomposition levels (e.g., [90]). The subsampling in this case is often referred to as multiscale (MS) after [103].

Previously, we considered the block-based CS paradigm primarily as an alternative to SRMs to mitigate computational burdens by limiting the CS measurement process to relatively small blocks. However, the drawback of BCS is typically a reduced quality of image reconstruction due to the fact that CS measurement generally works better the more global it is. We now consider an amelioration of the reconstruction quality of BCS while retaining its light computational burden and extremely fast execution. Specifically, as another primary contribution of this dissertation, we propose an MS algorithm that deploys BCS-SPL [82] in the domain of a wavelet transform; we note that this work was previously published as [48].

### **2.6.1 Wavelet-Domain BCS**

Ideally, the CS measurement operator should be “global” in the sense that the entire signal  $\mathbf{x}$  should contribute to each and every measurement taken in producing  $\mathbf{y}$  in (2.1). However, a block-diagonal structure as in (2.23) defeats such maximally holistic measurement. As a consequence, BCS-based techniques such as BCS-SPL and BCS-TV can be at a disadvantage in terms of reconstruction quality due to their reliance on a block-based measurement operator. We now consider a modification to the BCS-SPL algorithm—this variant of BCS-SPL is designed to improve its reconstruction-quality performance while maintaining its block-based measurement and corresponding fast reconstruction. Specifically, BCS-SPL is deployed within the wavelet domain of the image  $\mathbf{x}$  to provide multiscale measurement and reconstruction.

The measurement operator  $\Phi$  for MS-BCS-SPL is split into two components—a multiscale transform  $\Omega$  (e.g., a DWT) and a multiscale block-based measurement process  $\Phi'$  such that  $\Phi = \Phi'\Omega$ , and (2.1) becomes

$$\mathbf{y} = \Phi'\Omega\mathbf{x}. \quad (2.29)$$

Assume that  $\Omega$  produces  $L$  levels of wavelet decomposition; thus,  $\Phi'$  consists of  $L$  different block-based measurement operators, one for each level. That is, let the DWT of image  $\mathbf{x}$  be

$$\tilde{\mathbf{x}} = \Omega\mathbf{x}. \quad (2.30)$$

Subband  $s$  at level  $l$  of  $\tilde{\mathbf{x}}$  is then divided into  $B_l \times B_l$  blocks and measured using an appropriately-sized  $\Phi_l$  (note that  $l = L$  is the highest-resolution level). That is, suppose  $\tilde{\mathbf{x}}_{l,s,j}$  is a vector representing, in raster-scan fashion, block  $j$  of subband  $s$  at level  $l$ , with  $s \in \{H, V, D\}$ , and  $1 \leq l \leq L$ . Then,

$$\mathbf{y}_{l,s,j} = \Phi_l \tilde{\mathbf{x}}_{l,s,j}. \quad (2.31)$$

Since the different levels of wavelet decomposition have different importance to the final image reconstruction quality, we adjust the measurement process so as to yield a different subrate,  $S_l$ , at each level  $l$ . In all cases, we set the subrate of the DWT baseband to full measurement, i.e.,  $S_0 = 1$ . Then, we let the subrate for level  $l$  be

$$S_l = W_l S', \quad (2.32)$$

such that the overall subrate becomes

$$S = \frac{1}{4^L} S_0 + \sum_{l=1}^L \frac{3}{4^{L-l+1}} W_l S'. \quad (2.33)$$

Table 2.2

Wavelet-domain BCS subrates  $S_l$  at level  $l$  for target overall subrate  $S$  for a DWT with  $L = 3$  levels. In all cases, the baseband is given full measurement ( $S_0 = 1.0$ ).

$S$	<i>Level Subrates, <math>S_l</math></i>		
	$S_1$	$S_2$	$S_3$
0.1	1.0000	0.1600	0.0100
0.2	1.0000	0.5867	0.0367
0.3	1.0000	1.0000	0.0667
0.4	1.0000	1.0000	0.2000
0.5	1.0000	1.0000	0.3333

Given a target subrate  $S$  and a set of level weights  $W_l$ , one can easily solve (2.33) for  $S'$ , yielding a set of level subrates  $S_l$  via (2.32). However, this process will typically produce one or more  $S_l > 1$ . Thus, we modify the solution to enforce  $S_l \leq 1$  for all  $l$ . Specifically, after finding  $S'$  and  $S_1$  via (2.33) and (2.32), we check if  $S_1 > 1$ . If so, we set  $S_1 = 1$ , remove its corresponding term from the sum in (2.33), and then solve

$$S = \frac{1}{4^L} S_0 + \frac{3}{4^L} S_1 + \sum_{l=2}^L \frac{3}{4^{L-l+1}} W_l S' \quad (2.34)$$

for  $S'$ , again using (2.32) to redetermine  $S_l$  for  $l = 2, \dots, L$ . We repeat this process as needed to ensure that all  $S_l \leq 1$ .

For the experimental results to follow later, we use level weights,

$$W_l = 16^{L-l+1}, \quad (2.35)$$

which we have found to perform well in practice. The resulting level subrates  $S_l$  for various target subrates  $S$  for a DWT with  $L = 3$  levels are shown in Table 2.2.

### 2.6.2 Wavelet-Domain MS-BCS-SPL

The BCS-SPL reconstruction algorithm couples a full-image Wiener-filter smoothing process with a sparsity-enhancing thresholding process in the domain of some full-image sparsity transform  $\Psi$ . Interleaved between the smoothing and thresholding operations lie Landweber steps in the form of

$$\mathbf{x} \leftarrow \mathbf{x} + \Phi^T (\mathbf{y} - \Phi \mathbf{x}), \quad (2.36)$$

where  $\Phi$  is a measurement matrix. Figure 2.5 illustrates how the BCS-SPL reconstruction of Figure 2.2 is modified to accommodate the situation in which CS measurement takes place within a multiscale transform  $\Omega$  as in (2.29). In essence, the resulting MS-BCS-SPL reconstruction applies a Landweber step on each block of each subband in each decomposition level separately using the appropriate block-based  $\Phi_l$  for the current level  $l$ . As in the original BCS-SPL, Wiener filtering takes place in the spatial domain of the image, while some thresholding operator is applied in the domain of full-frame sparsity transform  $\Psi$  to promote sparsity.

```

function  $\tilde{\mathbf{x}} = \text{MS-BCS-SPL}(\mathbf{y}, \{\Phi_l, 1 \leq l \leq L\}, \Psi, \Omega)$ 
  for each level  $l$ 
    for each subband  $s \in \{H, V, D\}$ 
      for each block  $j$ 
         $\tilde{\mathbf{x}}_{l,s,j}^{(0)} = \Phi_l^T \mathbf{y}_{l,s,j}$ 
       $i = 0$ 
    do
       $\mathbf{x}^{(i)} = \Omega^{-1} \tilde{\mathbf{x}}^{(i)}$ 
       $\hat{\mathbf{x}}^{(i)} = \text{Wiener}(\mathbf{x}^{(i)})$ 
       $\hat{\tilde{\mathbf{x}}}^{(i)} = \Omega \hat{\mathbf{x}}^{(i)}$ 
      for each level  $l$ 
        for each subband  $s \in \{H, V, D\}$ 
          for each block  $j$ 
             $\hat{\tilde{\mathbf{x}}}_{l,s,j}^{(i)} = \hat{\tilde{\mathbf{x}}}_{l,s,j}^{(i)} + \Phi_l^T (\mathbf{y}_{l,s,j} - \Phi_l \hat{\tilde{\mathbf{x}}}_{l,s,j}^{(i)})$ 
           $\check{\tilde{\mathbf{x}}}^{(i)} = \Psi \Omega^{-1} \hat{\tilde{\mathbf{x}}}^{(i)}$ 
           $\check{\mathbf{x}}^{(i)} = \text{Threshold}(\check{\tilde{\mathbf{x}}}^{(i)})$ 
           $\bar{\tilde{\mathbf{x}}}^{(i)} = \Omega \Psi^{-1} \check{\mathbf{x}}^{(i)}$ 
        for each level  $l$ 
          for each subband  $s \in \{H, V, D\}$ 
            for each block  $j$ 
               $\tilde{\mathbf{x}}_{l,s,j}^{(i+1)} = \bar{\tilde{\mathbf{x}}}_{l,s,j}^{(i)} + \Phi_l^T (\mathbf{y}_{l,s,j} - \Phi_l \bar{\tilde{\mathbf{x}}}_{l,s,j}^{(i)})$ 
             $D^{(i+1)} = \|\tilde{\mathbf{x}}^{(i+1)} - \hat{\tilde{\mathbf{x}}}^{(i)}\|_2$ 
           $i = i + 1$ 
      until  $|D^{(i)} - D^{(i-1)}| < 10^{-4}$ 
     $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}^{(i)}$ 

```

Figure 2.5

MS-BCS-SPL reconstruction of a 2D image;  $\text{Wiener}(\cdot)$  is pixel-wise adaptive Wiener filtering using a neighborhood of  $3 \times 3$ , while  $\text{Threshold}(\cdot)$  is a thresholding process.

## 2.7 Other Approaches to CS Reconstruction of Images

After having considered block-based measurement and reconstruction for images in both the spatial and wavelet domains, we now overview several other prominent solutions for CS of images that have appeared in recent literature. Perhaps the most popular approach to the CS of images follows the MS paradigm [103] in which different decomposition levels in the wavelet domain are measured and reconstructed independently using a generic CS reconstruction. For example, in [5, 34, 58, 59, 72], this approach is applied using the generic reconstructions OMP [101], stagewise orthogonal matching pursuit (StOMP) [39], sparsity adaptive matching pursuits (SAMP) [34], CoSaMP [86], and Bayesian compressive sensing [5, 69]. As representative of such algorithms, we consider a multiscale variant of GPSR [46] as described in [90]. In essence, in this multiscale GPSR (MS-GPSR), GPSR reconstruction is applied independently to each DWT level; subrates in the individual levels are varied such that the baseband is retained in full, and successively higher-resolution decomposition levels feature a reduced subrate.

An example of a wavelet-domain CS measurement and reconstruction that is more sophisticated than the simple generic reconstructions like MS-GPSR involves the use of the well-known fact that wavelet coefficients inhabit a cross-scale tree structure in the transform domain (e.g., [94]). This tree-structured prior model has been incorporated into CS reconstruction both as an explicit coefficient structure [8] as well as in the form of a statistical model [58, 59]. In this latter statistical approach, called tree-structured wavelet compressed sensing (TSW-CS) [58, 59], a hierarchical Bayesian model is imposed on the coefficients across the DWT scales.



Fitting an *a priori* signal model to the CS reconstruction model is a frequently adopted strategy for the CS of images. In [72], a Gaussian scale mixture (GSM) [87] model is incorporated into several wavelet-domain CS reconstructions including reweighted  $\ell_1$  minimization [19], iteratively reweighted least squares [29], and iterative hard thresholding [13, 14]. Additionally, a spatial-domain piecewise autoregressive (PAR) model was incorporated into a spatial-domain CS reconstruction in [111]. This latter technique, called model-based adaptive recovery of compressive sensing (MARX) in [111], adopts the PAR model to better handle nonstationarity in the sparsity that images exhibit spatially.

Finally, there exist a number of algorithms for the general problem of inverse imaging that are often used in deconvolution or inpainting applications but can be applied to CS reconstruction as a special case. Such techniques include the algorithms of [10, 12, 110]. As representative of this general class of CS reconstruction, we consider the split augmented Lagrangian shrinkage algorithm (SALSA) of [1]. SALSA is a rather flexible solution to the CS reconstruction problem—when applied to CS, SALSA reconstructs from a full-image SRM-based measurement by enforcing either  $\ell_1$  sparsity in some transform domain (we use a DWT) or a minimum TV norm. We have found that DWT-based SALSA works better than its TV-based counterpart for the CS reconstruction of still images; additionally, we have found that the original SALSA implementation of [1] outperforms the later constrained variant of [2].

In the experimental results to follow, we compare several of the algorithms mentioned here to the BCS-SPL, MS-BCS-SPL, and TV algorithms discussed previously. Throughout our evaluations, we largely concern ourselves with algorithms that have implementations

readily available. Additionally, we restrict our attention to only those implementations that can handle images of a relatively large size (i.e.,  $512 \times 512$ ). This constraint effectively rules out algorithms which cannot accommodate SRM-based measurement operators (e.g., Bayesian compressive sensing [5, 69] and TSW-CS [58, 59], both of which directly access rows/columns of the measurement matrix) or which require an excessively long time to reconstruct a single image (e.g., MARX [111]). By “excessively long,” we mean more than several hours on a modern computer—given that we will be subsequently considering reconstruction of multiple frames for video, it is imperative that realistic reconstructions for video be based on still-image reconstructions that are significantly more expeditious.

## 2.8 Comparison of Various CS Techniques for Images

We now evaluate the performance of various algorithms described above on several gray-scale images of size  $512 \times 512$ . We compare the original BCS-SPL (Section 2.5.1 and [82]) with spatial-domain BCS, MS-BCS-SPL (Section 2.6.2 and [48]) with wavelet-domain BCS, TV reconstruction [17], MS-GPSR [90], and SALSA [1]. Both MS-BCS-SPL and BCS-SPL use a DDWT [73] as the sparsity transform  $\Psi$  with bivariate shrinkage [26] applied within the DDWT domain to enforce sparsity as described in [82]. MS-BCS-SPL uses a 3-level DWT with the popular 9/7 biorthogonal wavelets [3] as the measurement-domain transform  $\Omega$ . At decomposition level  $l$  of  $\Omega$ , blocks of size  $B_l \times B_l$  are individually measured in the DWT domain using the scrambled block-DCT SRM measurement operator of [35]; we use blocks of sizes  $B_l = 16, 32$ , and  $64$  for decomposition levels  $l = 1, 2$ , and  $3$ , respectively ( $l = 3$  is the highest-resolution level). On the other hand, BCS-SPL

uses  $B \times B$  block-based measurement applied directly on the image in its ambient domain; here,  $B = 32$ . TV and SALSA use the scrambled block-Hadamard SRM of [51] to provide a fast whole-image, spatial-domain CS measurement; additionally, SALSA employs a 9/7 biorthogonal DWT as a sparsity transform. Finally, MS-GPSR is implemented similarly to MS-BCS-SPL—GPSR reconstruction is applied independently to each DWT level using the same  $\Omega$  as MS-BCS-SPL; subrates in the individual levels follow Table 2.2 with measurement using a scrambled block-DCT SRM applied to the entire DWT level. We use our implementation<sup>8</sup> of BCS-SPL and MS-BCS-SPL,  $\ell_1$ -MAGIC<sup>9</sup> for TV, and the GPSR<sup>10</sup> and SALSA<sup>11</sup> implementations from their respective authors. We note that the results presented here were previously reported in [49].

The reconstruction performance of the various algorithms under consideration is presented in Table 2.3. In most cases, the wavelet-domain measurement and MS reconstruction of MS-BCS-SPL provides a substantial gain in reconstruction quality over the spatial-domain measurement of BCS-SPL, generally on the order of a 1- to 3-dB increase in PSNR. Additionally, MS-BCS-SPL outperforms TV reconstruction in most instances despite the fact that TV has the advantage of full-image measurement; the gains of MS-BCS-SPL over TV are particularly significant at the lowest subrates. MS-BCS-SPL also generally outperforms MS-GPSR even though the latter globally measures each resolution level. The primary exception is the “Barbara” image—although MS-BCS-SPL outperforms TV at the lowest subrates, MS-GPSR is slightly better. However, TV dominates the performance

---

<sup>8</sup><http://www.ece.msstate.edu/~fowler/BCSSPL/>

<sup>9</sup><http://www.l1-magic.org>

<sup>10</sup><http://www.lx.it.pt/~mtf/GPSR/>

<sup>11</sup><http://cascais.lx.it.pt/~mafonso/salsa.html>

Table 2.3

Reconstruction PSNR in dB; images are  $512 \times 512$ .

<i>Algorithm</i>	<i>Subrate</i>				
	0.1	0.2	0.3	0.4	0.5
Lenna					
MS-BCS-SPL	<b>31.6</b>	<b>34.7</b>	<b>36.7</b>	<b>37.9</b>	<b>39.0</b>
BCS-SPL	28.0	31.6	33.7	35.4	36.9
TV	29.9	32.9	35.0	36.8	38.4
MS-GPSR	30.3	33.6	35.2	36.3	37.8
SALSA	23.9	28.5	31.6	34.0	36.0
Barbara					
MS-BCS-SPL	23.8	25.1	26.1	27.4	28.8
BCS-SPL	22.4	23.8	25.4	27.0	28.7
TV	23.0	24.5	<b>26.3</b>	<b>28.4</b>	<b>30.8</b>
MS-GPSR	<b>24.0</b>	<b>25.3</b>	26.1	27.5	29.6
SALSA	19.7	22.7	25.1	27.7	30.4
Peppers					
MS-BCS-SPL	<b>31.1</b>	<b>34.2</b>	<b>35.7</b>	<b>36.8</b>	<b>37.7</b>
BCS-SPL	29.0	32.1	33.8	35.2	36.4
TV	30.4	33.1	34.7	35.9	37.0
MS-GPSR	29.3	31.9	33.1	34.2	35.8
SALSA	23.3	28.2	31.2	33.3	35.0
Mandrill					
MS-BCS-SPL	21.4	<b>23.0</b>	<b>24.6</b>	<b>25.5</b>	<b>26.5</b>
BCS-SPL	20.5	21.8	22.9	23.9	25.1
TV	20.5	22.0	23.4	24.9	<b>26.5</b>
MS-GPSR	<b>21.5</b>	22.9	24.3	25.1	26.3
SALSA	16.6	19.6	21.1	22.5	24.2
Goldhill					
MS-BCS-SPL	<b>29.0</b>	<b>31.1</b>	<b>32.8</b>	<b>33.7</b>	34.7
BCS-SPL	27.1	29.1	30.5	31.8	33.1
TV	27.5	29.9	31.6	33.2	<b>34.8</b>
MS-GPSR	28.5	30.4	32.2	33.0	34.1
SALSA	22.9	26.0	28.2	30.2	32.0

Table 2.4

Reconstruction time for the  $512 \times 512$  “Lenna” image at subrate of 0.3.

<i>Algorithm</i>	<i>Time (sec.)</i>
BCS-SPL	30
MS-BCS-SPL	46
SALSA	111
MS-GPSR	1,173
TV	6,584

comparison for “Barbara” at the higher subrates. Figure 2.6 depicts typical reconstruction results for the “Lenna” image using a subrate  $S = 0.1$ ; these figures show only a detailed portion from the center of the image.

As can be seen in Table 2.4, in terms of execution times, reconstruction with MS-BCS-SPL is only slightly slower than BCS-SPL, each running for about half a minute on a dual-core 2.8-GHz machine. On the other hand, SALSA is somewhat slower, while the execution times of both MS-GPSR and TV are some two orders of magnitude longer, with TV requiring nearly two hours to reconstruct a single image despite the use of a fast SRM measurement operator.

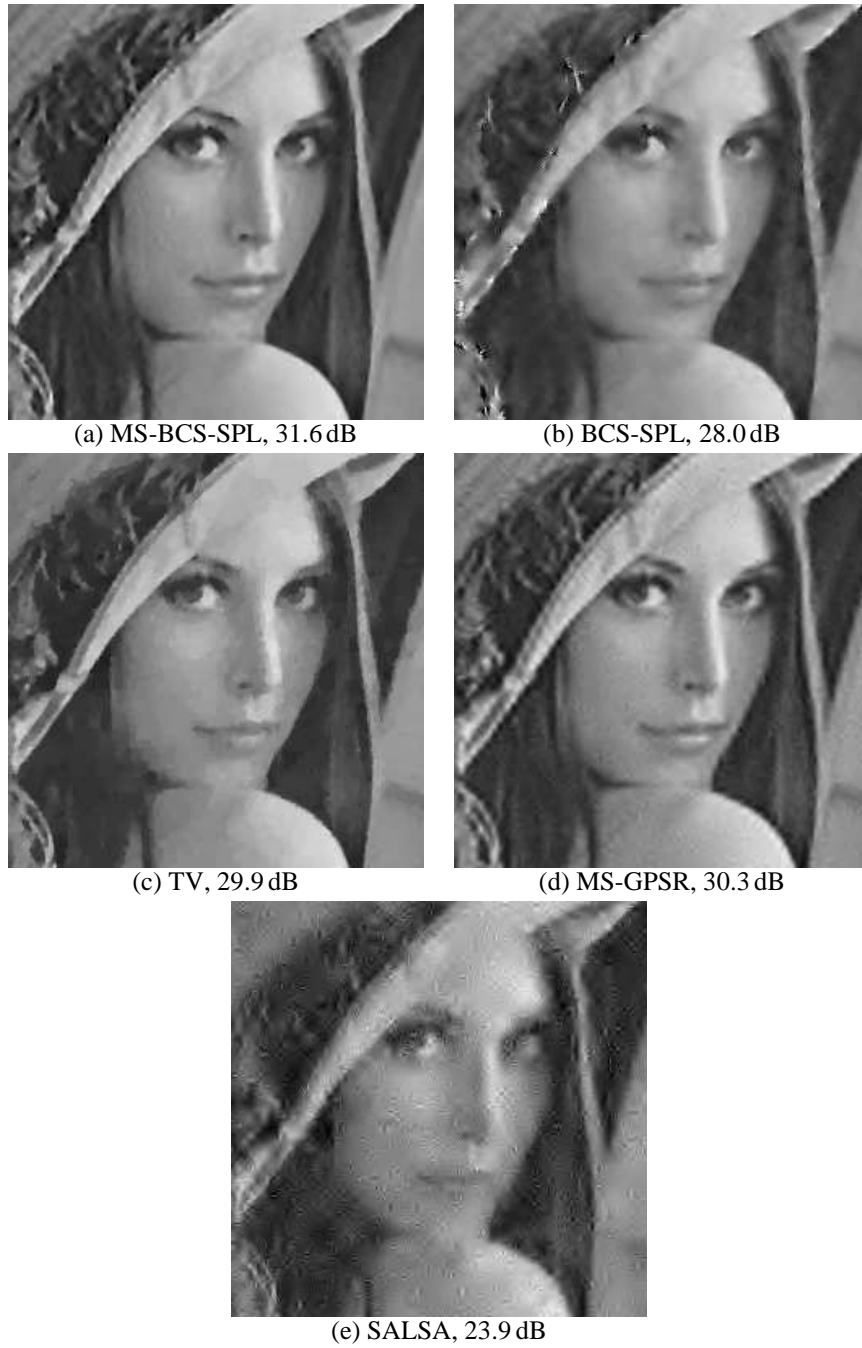


Figure 2.6

Reconstructions of the  $512 \times 512$  “Lenna” image (shown in detail) for a subrate of  $S = 0.1$ .

## 2.9 Remarks

In applying the CS paradigm to still images, two primary challenges must be addressed. First, the CS reconstruction should be specifically tailored to the fact that the underlying signal is an image, and, second, the measurement and reconstruction processes must accommodate the large-sized signals that accompany multidimensional data without imposing large computational or memory burdens. Above, we have discussed several solutions for the first issue, including the imposition of smoothness, directional sparsity transforms, as well as CS measurement in the wavelet domain. For the second issue, we have considered the use of both SRMs as well as block-based measurement operators in both the measurement and reconstruction processes.

Overall, the MS-BCS-SPL algorithm (Section 2.6.2 and [48]) effectively addresses both issues—MS-BCS-SPL retains the fast execution speed associated with block-based measurement while rivaling the quality of CS reconstructions such as TV that employ full-image measurement. However, there exist practical-implementation issues for any CS technique employing wavelet-domain measurement, including the MS-BCS-SPL and MS-GPSR [90] approaches considered here as well as other prominent CS reconstructions such as TSW-CS [58, 59].

Specifically, the general advantages of BCS are a reduced computational complexity in reconstruction as well as a greatly simplified measurement-operator implementation in both the reconstruction as well as sensing processes. A multiscale BCS in the wavelet domain like that used in MS-BCS-SPL retains these advantages for reconstruction; however, the decomposition of the measurement process as  $\Phi = \Phi' \Omega$  entails that the transform  $\Omega$  (a

dense matrix) wrecks the block-diagonal structure of  $\Phi'$ . As a consequence, the resulting  $\Phi$  becomes dense and thus a challenge to implement within a CS sensing device. As a consequence, wavelet-domain measurement defeats the computation and memory advantages of BCS in the sensing device, although these advantages can still be exploited on the reconstruction side of the system. Furthermore, wavelet-domain measurement with most popular wavelets will require the measurement matrix  $\Phi$  to be real-valued even if the underlying  $\Phi'$  is binary or Rademacher, thereby requiring duty cycling in a single-pixel camera and further complicating the acquisition process. However, this implementation challenge can be solved using different sensing device in the single pixel camera. Instead of using the original DMD which only reflects two different directions, the use of a different type of DMD which can produce 1024 (10 bit) grayscale values [60] enables the wavelet domain sensing because Haar wavelet and a pseudorandom Rademacher  $\pm 1$  measurement basis can be implemented in the form of (2.15) [4].

In next chapter, we turn our attention from the reconstruction of a single still image and toward applications dealing with multiple images, namely, video sequences. In this application, we will find ourselves reconstructing a set of multiple images, a task that we could, of course, accomplish by reconstructing each image independently. However, we shall see that significant performance gains can be obtained through reconstructing the multiple images while capitalizing on any correlation that exists across the image set. In any event, we will require a very fast still-image recovery as the foundation of CS reconstruction for this more complicated, multiple-image data in the remainder of this work, since the still-image reconstruction will be employed numerous times. Due to its exceedingly fast computa-



tion, we therefore employ BCS-SPL as the still-image reconstruction engine upon which mechanisms for the exploitation of cross-image correlation are built.

## CHAPTER 3

### BLOCK-BASED COMPRESSED SENSING FOR VIDEO

In much of the literature in the area of compressed sensing (CS), the primary goal has largely been the blind reconstruction of CS-acquired signals. That is, for the most part, many CS reconstruction strategies are oblivious to the structure of the signal being recovered beyond a general assumption of sparsity in some transform basis. Recently, however, several reconstruction approaches have focused on situations in which additional information about the signal content is available to aid signal reconstruction; sometimes this additional information is called “side information.” Of specific interest here is the situation in which one or more predictions of the signal to be recovered are available to the CS reconstruction process. Video sequences are one form of data in particular in which predictions are commonly used in various forms of processing. Specifically, it is typical in video processing that one or more reference frames are used to make predictions of some current frame such that the resulting residual frame has dramatically lowered signal energy leading to more efficient representation and processing. This paradigm is, in fact, fundamental to the traditional source coding of video and an essential part of all modern video standards such as MPEG [61, 62] and H.264/AVC [64].

We now consider the CS reconstruction of video sequences in which frame-to-frame predictions are used to aid the CS-reconstruction process. As a primary contribution of

this dissertation, we investigate CS reconstruction on the prediction residual which is, in most cases, significantly more *compressible*, in the sense of (2.4), than the original frame; such prediction thus results in a higher-quality CS reconstruction. Key to this prediction-driven residual reconstruction is the use of motion estimation (ME) and motion compensation (MC) such that the frame-to-frame predictions compensate for object motion between frames. Such use of ME/MC derives from traditional video-coding algorithms which make extensive use of sophisticated MC strategies. We note that the work that is presented in this chapter was previously published as [49, 83].

### 3.1 Block-Based Acquisition of Video

The straightforward implementation of CS on video would involve, as for images, the vectorization of a 3D group of frames into a single 1D vector. However, the computation and memory issues associated with this vectorization approach for 2D images are exacerbated with the even greater dimensionality present in video data. Additionally, a global sampling simultaneously across the spatial and temporal extent of a group of frames is likely to be impractical [40]; this is due to the fact that a CS device must access the entirety of a video signal in both time and space in order to calculate the requisite random projections for CS. Consequently, we focus on the situation in which video frames are acquired independently in a 2D fashion, e.g., by applying a suitable image acquisition (e.g., single-pixel camera) in a frame-by-frame fashion as in [106, 107]. Again, to cut computation and memory, we focus on a block-based image measurement applied frame by frame.

### 3.2 Straightforward CS Reconstruction for Video

We now consider several strategies for the CS reconstruction of video frames acquired the frame-by-frame strategy described above. These include both intraframe and 3D versions of BCS-SPL.

Given a frame-by-frame acquisition, the most straightforward reconstruction would be to reconstruct the individual frames independently using the BCS-SPL procedure introduced previously. However, such an intraframe BCS-SPL reconstruction ignores the fact that consecutive video frames are usually highly correlated. Nonetheless, intraframe BCS-SPL serves as a baseline against which to compare other techniques that do attempt to exploit such temporal correlation.

Another straightforward method of CS reconstruction for video that does make an effort to exploit temporal correlation is to treat the video frames as a 3D “volume” and reconstruct the 3D video volume by applying a suitable CS reconstruction algorithm using a 3D transform; this was done, for example, in [106, 107]. Although such 3D reconstruction could be applied across a video volume as a whole (as in [106, 107]), for computation and memory issues (which are likely to be substantial given the increased dimensionality), we consider a 3D version of the BCS-SPL algorithm, essentially extending BCS-SPL reconstruction for a single still image into three dimensions, with the video volume being partitioned into smaller, 3D cubes.

Specifically, let us consider a group of  $P$  consecutive frames from a video sequence which we call a group of pictures (GOP). Again, for video acquisition, the individual video frames are acquired using a 2D block-based measurement applied frame by frame. For

reconstruction, the GOP is partitioned into  $B \times B \times P$  cubes; i.e., 3D blocks which have a spatial size of  $B \times B$  and a temporal size of  $P$ . The BCS-SPL reconstruction then uses a 3D block-based transform operator; e.g., a 3D discrete cosine transform (DCT) of size  $B \times B \times P$ . For thresholding, we employ a variation on (2.27) wherein the median operator is replaced by the mean (i.e.,  $\sigma^{(i)} = \text{mean}(|\check{\mathbf{x}}^{(i)}|)/0.6745$ ); we have observed empirically that this mean-based estimator yields faster convergence in the Landweber operation for video. We refer to the resulting technique as 3D-BCS-SPL reconstruction.

It was argued in [107] that, in effect, CS reconstruction using a 3D transform—such as the 3D-BCS-SPL discussed here—attempts to exploit the frame-to-frame correlation that exists within a video GOP through the joint sparsity that occurs in the 3D transform. This is in contrast to intraframe BCS-SPL reconstruction which makes no attempt at all to exploit frame-to-frame correlation. However, neither of these two approaches takes full advantage of the temporal correlation that exists in the video sequence due to the frame-to-frame motion of objects. For this, some form of ME/MC must be incorporated into the reconstruction; we consider such ME/MC-based reconstruction next.

### 3.3 The Motion-Compensated BCS-SPL Algorithm

In video coding, knowledge of object motion is used to make interframe predictions which are, in turn, used to drive an efficient coding of prediction residuals. As a result, ME/MC is a widely used and crucial component to traditional video-coding systems. We now incorporate this ME/MC framework into the reconstruction process of BCS-SPL. This implies that ME/MC resides at the reconstruction, or “decoder,” side of a BCS-SPL sys-

tem for video, rather than at the sensing, or “encoder,” side as is the case in traditional video coding. As a primary contribution of this dissertation, we propose an extension of BCS-SPL for video that uses the same simple measurement as was used for 2D images previously—block-based random CS measurements are applied frame by frame. The resulting motion-compensated version of BCS-SPL, or motion-compensated BCS-SPL (MC-BCS-SPL), consists of several main components: residual reconstruction, in which BCS-SPL is applied to an MC residual; multihypothesis initialization, in which intraframe BCS-SPL is used to initialize the MC-BCS-SPL reconstruction; and forward/backward MC, in which multiple reconstruction passes are performed in multiple directions across the GOP. These components are explored below. We note that we previously introduced in MC-BCS-SPL in [49, 83].

### 3.3.1 Residual Reconstruction

Suppose we have two consecutive frames, the current frame,  $\mathbf{x}$ , and the reference frame,  $\mathbf{x}_{\text{ref}}$ . We assume that the reference frame has been previously reconstructed, while we have only random measurements,  $\mathbf{y}$ , of the current frame, using block-based measurement. As a first step, we could reconstruct an approximation,  $\hat{\mathbf{x}}$ , to current image  $\mathbf{x}$  by simply applying the BCS-SPL image reconstruction to  $\mathbf{y}$  (we consider a more sophisticated estimate of  $\hat{\mathbf{x}}$  later). However, such intraframe reconstruction does not capitalize on our knowledge of reference frame  $\mathbf{x}_{\text{ref}}$ . Yet, at this point, we have approximations to the current frame as well as to the reference frame—we can perform ME on these two frames to estimate a motion field describing motion of objects between the two frames. Such ME

could be performed in a variety of ways; for simplicity we consider full-search, block-based ME as is commonly used in video coding. Consequently, the approximation to the current frame,  $\hat{\mathbf{x}}$ , is partitioned into blocks whose motion from the reference frame is indicated by a field of motion vectors. This permits the production of a motion-compensated frame,  $\hat{\mathbf{x}}_{\text{mc}}$ , that forms a prediction of  $\hat{\mathbf{x}}$  and thus also the still-unknown true current frame  $\mathbf{x}$ . Using this prediction, we can form a “projection-domain” residual by simply applying the known block-based random measurement operator to the motion-compensated frame; i.e., for each block  $j$ :

$$\mathbf{y}_{r_j} = \mathbf{y}_j - \Phi_B \hat{\mathbf{x}}_{\text{mc}_j}. \quad (3.1)$$

It is clear that  $\mathbf{y}_{r_j}$  is the random projection of the residual,  $\mathbf{x}_{r_j}$ , between our motion-compensated prediction  $\hat{\mathbf{x}}_{\text{mc}_j}$  and the original and still-unknown block  $\mathbf{x}_j$ ; i.e.,

$$\mathbf{y}_{r_j} = \mathbf{y}_j - \Phi_B \hat{\mathbf{x}}_{\text{mc}_j} = \Phi_B (\mathbf{x}_j - \hat{\mathbf{x}}_{\text{mc}_j}) = \Phi_B \mathbf{x}_{r_j}. \quad (3.2)$$

If the MC process is accurate, the residual frame  $\mathbf{x}_r$  should be more compressible—in the sense of (2.4)—than the original image  $\mathbf{x}$ . This is demonstrated empirically for a video sequence in Figure 3.1 wherein it is seen that, in reference to (2.4), the transform-coefficient magnitudes decay more quickly for the residual frame  $\mathbf{x}_r$  than for the original frame  $\mathbf{x}$ . In such a case, CS reconstruction will thereby be more effective at recovering the residual  $\mathbf{x}_r$  from  $\mathbf{y}_r$  than it is at recovering  $\mathbf{x}$  from  $\mathbf{y}$ . Let  $\hat{\mathbf{x}}_r$  be such a BCS-SPL reconstruction from  $\mathbf{y}_r$ ; consequently, we can form a new approximation to  $\mathbf{x}$  as

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{mc}} + \hat{\mathbf{x}}_r. \quad (3.3)$$

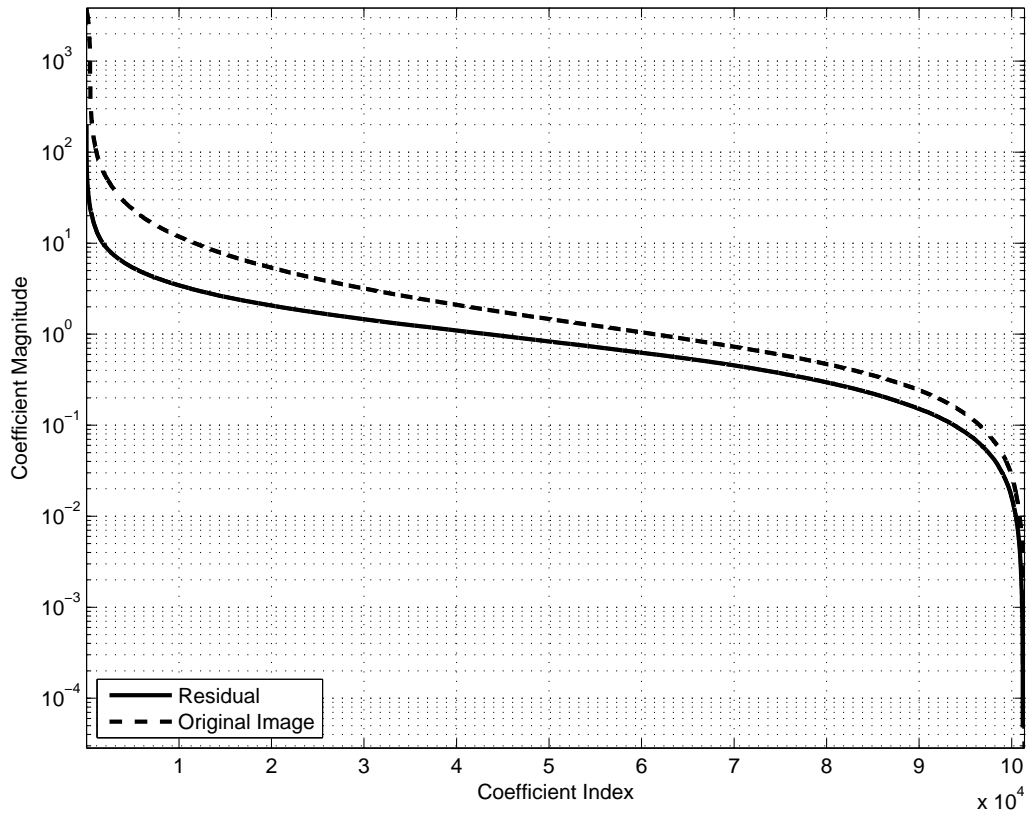


Figure 3.1

Decay of the magnitudes of the transform coefficients for frame 1 of the “Foreman” video sequence as compared to that of the motion-compensated residual between frames 1 and 0. The transform is a 4-level biorthogonal 9/7 DWT.



```

function  $\hat{\mathbf{x}} = \text{MC-BCS-SPL}(\mathbf{y}, \Phi_B, \Psi, \mathbf{x}_{\text{ref}})$ 
   $\hat{\mathbf{x}} = \text{Initialize}(\mathbf{y}, \Phi_B, \Psi, \mathbf{x}_{\text{ref}})$ 
   $i = 0$ 
  while  $i < \text{MAX\_ITERATION}$ 
     $\hat{\mathbf{x}}_{\text{mc}} = \text{MotionCompensation}(\hat{\mathbf{x}}, \mathbf{x}_{\text{ref}})$ 
    for each block  $j$ 
       $\mathbf{y}_{\text{mc}_j} = \Phi_B \hat{\mathbf{x}}_{\text{mc}_j}$ 
     $\mathbf{y}_r = \mathbf{y} - \mathbf{y}_{\text{mc}}$ 
     $\hat{\mathbf{x}}_r = \text{BCS-SPL}(\mathbf{y}_r, \Phi_B, \Psi)$ 
     $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{mc}} + \hat{\mathbf{x}}_r$ 
     $i = i + 1$ 
  end while

```

Figure 3.2

MC-BCS-SPL reconstruction of the current frame from a single reference frame.

We now have a new approximation to the current frame that is of better quality than the initial approximation that we created from a direct BCS-SPL reconstruction from  $\mathbf{y}$ . Consequently, we will be able to produce a more accurate motion-vector field from ME applied to this new  $\hat{\mathbf{x}}$  and  $\mathbf{x}_{\text{ref}}$ , and further enhancement can be expected by iteratively repeating the above process. The resulting MC-BCS-SPL algorithm is summarized in Figure 3.2, where  $\text{BCS-SPL}(\cdot)$  is the BCS-SPL recovery from Figure 2.2, and  $\text{Initialize}(\cdot)$  is described in Figure 3.3.

### 3.3.2 Multihypothesis Initialization

We now return to the issue of producing an initial estimate,  $\hat{\mathbf{x}}$ , of the current frame. Let us assume two frames again, reference frame,  $\mathbf{x}_{\text{ref}}$ , and random measurement,  $\mathbf{y}$ , of the current frame,  $\mathbf{x}$ . As mentioned above, one possible way to obtain an initial guess for the current frame is to apply BCS-SPL; let this estimate be  $\hat{\mathbf{x}}'$ . An alternative approach would be a residual reconstruction of the frame-to-frame difference with the motion field set to zero. The former approximation might be suitable for a dynamic sequence with high-motion content. On the other hand, some natural video sequences are relatively stationary, such as newscasts and surveillance video whose motion vectors are mostly zero or close to zero; in these cases, a residual reconstruction with a zero motion field might be more appropriate as an initial reconstruction. In this zero-motion case, the residual reconstruction (3.1) can be rewritten as

$$\mathbf{y}_{\text{r}_j} = \mathbf{y}_j - \Phi_B \mathbf{x}_{\text{ref}_j}, \quad (3.4)$$

leading to an alternate estimate,  $\hat{\mathbf{x}}''$ , of the current frame.

Both of these approaches—BCS-SPL reconstruction as well as zero-motion residual reconstruction—provide a guess for the current frame, but neither is likely to be exactly correct, and each might be somewhat different than the true current frame. As a consequence, as the final initial frame, we average these two approximations as the initial guess for the MC-BCS-SPL reconstruction. This process, which is depicted algorithmically in Figure 3.3, is similar to multihypothesis MC (e.g., [53, 95]) in traditional video coding, so we refer to it as multihypothesis initialization. We have observed empirically over a variety of video sequences that this multihypothesis initialization outperforms either of the

```

function  $\hat{\mathbf{x}} = \text{Initialize}(\mathbf{y}, \Phi_B, \Psi, \mathbf{x}_{\text{ref}})$ 
 $\hat{\mathbf{x}}' = \text{BCS-SPL}(\mathbf{y}, \Phi_B, \Psi)$ 
for each block  $j$ 
     $\mathbf{y}_{r_j} = \mathbf{y}_j - \Phi_B \mathbf{x}_{\text{ref}_j}$ 
 $\hat{\mathbf{x}}'' = \text{BCS-SPL}(\mathbf{y}_r, \Phi_B, \Psi) + \mathbf{x}_{\text{ref}}$ 
 $\hat{\mathbf{x}} = \frac{1}{2} [\hat{\mathbf{x}}' + \hat{\mathbf{x}}'']$ 

```

Figure 3.3

Multihypothesis initialization in MC-BCS-SPL reconstruction.

single hypotheses,  $\hat{\mathbf{x}}'$  or  $\hat{\mathbf{x}}''$ , used alone. Sample results are illustrated in Figure 3.4. Results are for frame 1 of the “Coastguard” sequence with frame 0 as the reference frame  $\mathbf{x}_{\text{ref}}$ ; the reference frame is reconstructed independently using BCS-SPL with a subrate of 0.7. Performance is shown for MC-BCS-SPL reconstruction of the current frame using the designated initialization (see Figure 3.3) for the current frame: intraframe initialization ( $\hat{\mathbf{x}}'$ ), zero-motion residual initialization ( $\hat{\mathbf{x}}''$ ), or multihypothesis initialization ( $[\hat{\mathbf{x}}' + \hat{\mathbf{x}}''] / 2$ ).

According to the empirical results, in the absence of significant motion, multihypothesis initialization permits the initial reconstruction to exploit the reference frame in the case that it has higher quality than the intraframe reconstruction of the current frame (which is, in fact, likely, due to the forward-backward reconstruction process which we discuss next).

### 3.3.3 Forward/Backward Motion Compensation

Thus far, we have considered the reconstruction of two consecutive frames of a video sequence. We now turn our attention to the more realistic problem of applying MC-BCS-

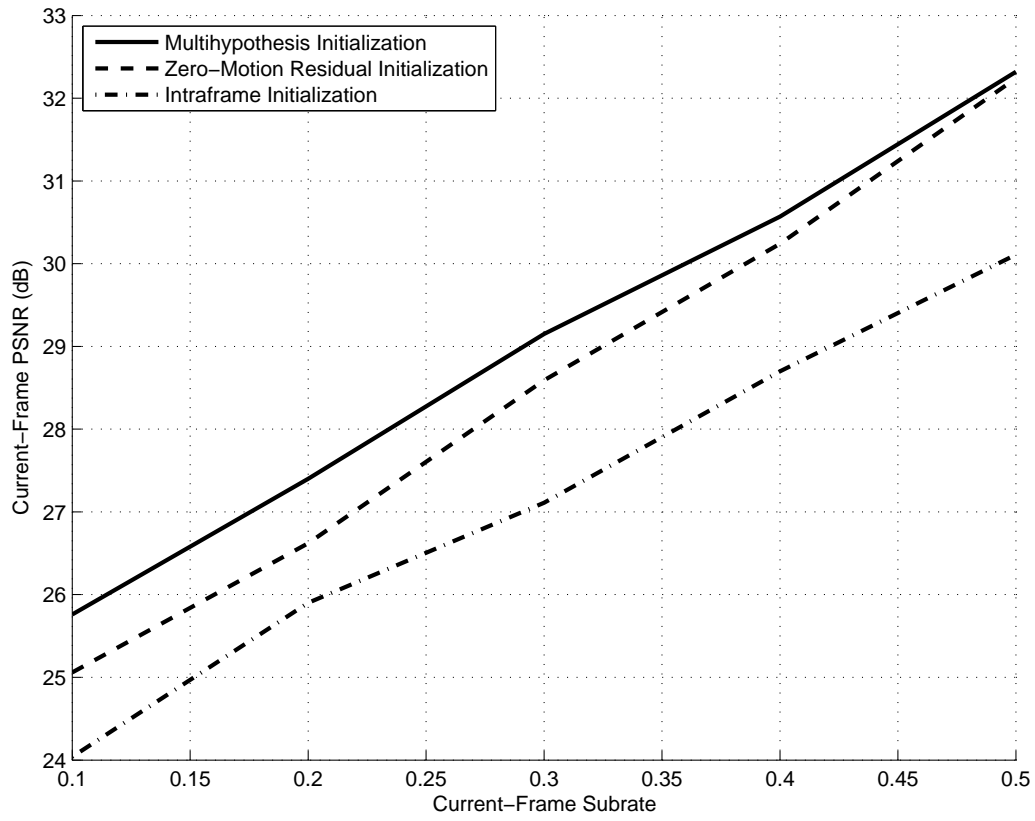


Figure 3.4

Various strategies for the initial estimate of the current frame in MC-BCS-SPL.

SPL to multiple frames. Specifically, let us consider a GOP of  $P$  consecutive frames from a video sequence, consisting of one “key frame” (the first frame) followed by  $P - 1$  “non-key frames.”

In the MC-BCS-SPL setting, we have block-based random measurements of each of the frames of the GOP; i.e.,  $\mathbf{y}_p = \Phi_p \mathbf{x}_p$  for  $0 \leq p \leq P - 1$ , where  $p$  is the frame number.  $\Phi_p$  is the random block-based measurement operator for frame  $p$ ; we assume that this measurement operator has subrate of  $S_p = M_p/N$  in that it reduces the  $N$ -dimensional  $\mathbf{x}_p$  image signal to an  $M_p$ -dimensional measurement signal  $\mathbf{y}_p$ . We will focus our attention on two cases: 1) all frames are acquired with the same subrate, and 2) the key frame,  $\mathbf{x}_0$ , is acquired at a relatively high subrate, while all the non-key frames are at an identical lower subrate.

Direct, intraframe BCS-SPL reconstruction from  $\mathbf{y}_0$  produces  $\hat{\mathbf{x}}_0$ , a reconstruction of the key frame which is then used as the reference frame for the MC-BCS-SPL reconstruction of the second frame of the GOP. This MC-BCS-SPL process produces  $\hat{\mathbf{x}}_1$  from  $\mathbf{y}_1$  and  $\hat{\mathbf{x}}_0$ ;  $\hat{\mathbf{x}}_1$  is then in turn used as the reference frame in MC-BCS-SPL reconstruction of  $\hat{\mathbf{x}}_2$  from  $\mathbf{y}_2$ . This process continues through in this manner to reconstruct the remaining frames of the GOP.

We note, however, that, since each frame is used as a reference for a subsequent frame, we have observed a successively lower reconstruction quality as we progress through the non-key frames. That is, we have observed reconstruction quality to deteriorate with increasing frame number  $p$ , such that the quality of  $\hat{\mathbf{x}}_{P-1}$ , the reconstructed last frame of the GOP, will be much less than that of  $\hat{\mathbf{x}}_0$ , the first frame of the GOP. In [83], it is proposed

to combat this quality deterioration by performing another iteration of MC-BCS-SPL reconstruction on the GOP, but this time running in the reverse temporal direction, from frame  $P - 1$  to frame 0.

Specifically, we assume that there exists another GOP following the current GOP, with  $\mathbf{x}_P$  being the first frame of that GOP (here, we abuse our notation by indexing this frame from the next GOP relative to the current GOP). Direct BCS-SPL reconstruction from  $\mathbf{y}_P$  will produce a reconstruction  $\hat{\mathbf{x}}_P$ ; this reconstruction can be used as the reference frame for the MC-BCS-SPL reconstruction of  $\mathbf{y}_{P-1}$  to produce  $\hat{\mathbf{x}}_{P-1}$ . We could continue MC-BCS-SPL reconstruction in this reverse temporal direction until we reached the start of the current GOP.

Running MC-BCS-SPL reconstruction on the current GOP in both the forward and backward temporal directions would yield two reconstructions for each non-key frame of the GOP. Our experimental observations have revealed that MC-BCS-SPL in the forward temporal direction yields the higher-quality reconstruction for the first half of the GOP, while MC-BCS-SPL in the backward temporal direction yields the better quality for the last half of the GOP. Assume that the GOP consists of an even number of frames. Thus, in [83], it is proposed to use forward MC-BCS-SPL to reconstruct frames  $\hat{\mathbf{x}}_p$  for  $1 \leq p \leq \frac{P}{2} - 1$ , the first half of the non-key frames, and backward MC-BCS-SPL (starting from the key frame of the next GOP) to reconstruct frames  $\hat{\mathbf{x}}_p$  for  $\frac{P}{2} + 1 \leq p \leq P - 1$ . To reconstruct the center frame,  $\hat{\mathbf{x}}_{P/2}$ , we slightly modify the MC-BCS-SPL procedure of Figure 3.2 to incorporate bidirectional MC prediction. Specifically, MC-BCS-SPL iterates first using  $\hat{\mathbf{x}}_{P/2-1}$  (the forward reference frame) as the reference frame for the MotionCompensation( $\cdot$ ) operator,

and then iterates using  $\hat{\mathbf{x}}_{P/2+1}$  (the backward reference frame). This process is repeated for an additional set of both forward and backward iterations to yield the final reconstruction of the center frame. Figure 3.5 illustrates this forward/backward MC-BCS-SPL reconstruction process, where black frames are direct (intraframe) BCS-SPL reconstruction of key frames, and white frames are forward/ backward MC-BCS-SPL reconstruction of non-key frames (forward MC-BCS-SPL for frames 1–3, backward MC-BCS-SPL for frames 5–7), and gray frame is bidirectional MC-BCS-SPL reconstruction of center non-key frame.

Because center frame  $\hat{\mathbf{x}}_{P/2}$  is recovered using both forward and backward reconstruction, it turns out to usually have a relatively high reconstruction quality. As a second and final stage of reconstruction, we update the frames between the center frame ( $\hat{\mathbf{x}}_{P/2}$ ) and the key frames ( $\hat{\mathbf{x}}_0$  and  $\hat{\mathbf{x}}_P$ ) by performing the same bidirectional reconstruction as was used on the center frame. A complete description of the two phases of reconstruction is given in Table 3.1. In this table, Phase 1 is forward/backward reconstruction from Figure 3.5, while Phase 2 is “enhancement” reconstruction from the center frame toward the two key frames in reverse direction.

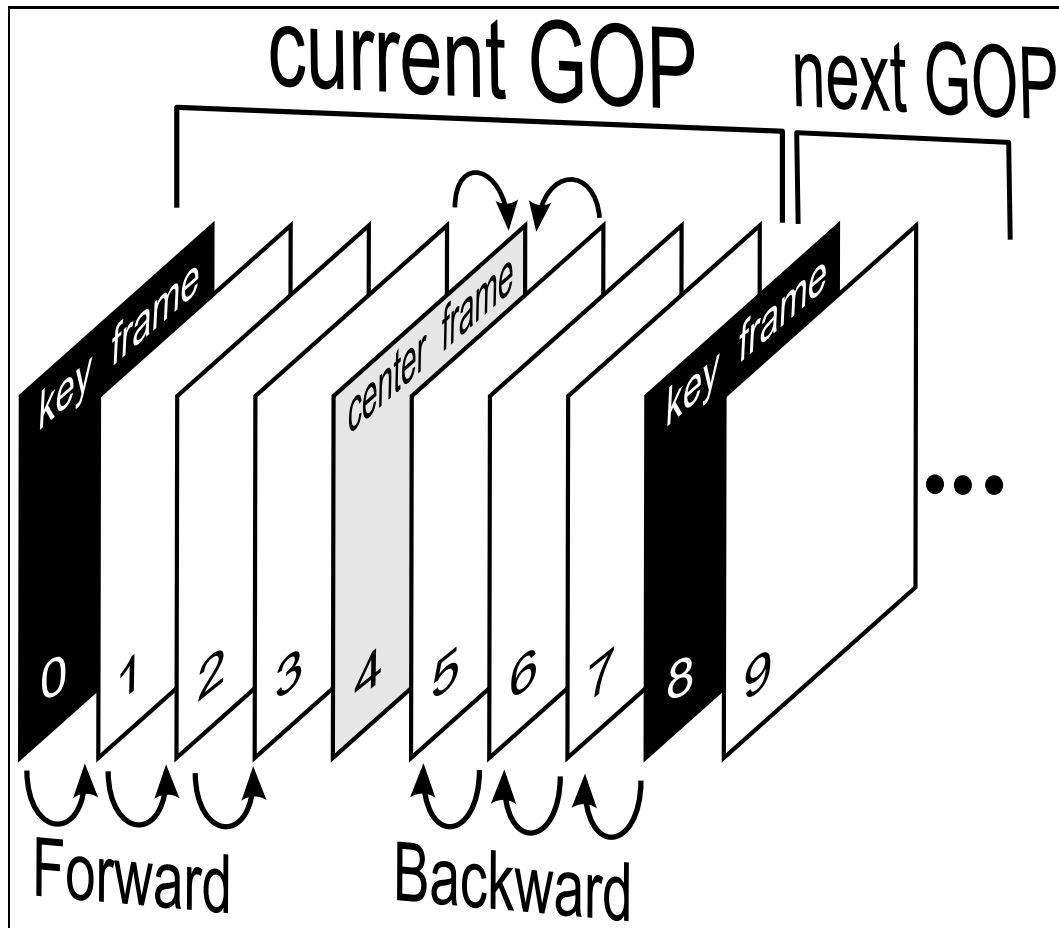


Figure 3.5

MC-BCS-SPL reconstruction applied to a GOP of size  $P = 8$  frames.



Table 3.1

Two-phase multiframe reconstruction for a GOP of size  $P = 8$ .

	Current frame	Reference frame
Phase 1	1	0
	2	1
	3	2
	8	7
	7	6
	6	5
	4	3, 5*
Phase 2	3	4
	5	4
	2	1, 3*
	1	2
	3	2
	6	5, 7*
	5	6
	7	6

\*bidirectional reconstruction

### 3.4 Experimental Observations

We now examine the performance of MC-BCS-SPL reconstruction relative to its simple intraframe and 3D volumetric variants to demonstrate that significant gain results from the explicit exploitation of motion information within the CS reconstruction of video. We use the common video sequences “Coastguard” (296 frames), “Football” (120 frames), “Foreman” (296 frames), “Hall Monitor” (88 frames), “Mobile” (296 frames), “Mother and Daughter” (296 frames), “Stefan” (296 frames), and “Susie” (72 frames). These sequences have grayscale CIF frames of size  $352 \times 240$  or  $352 \times 288$ . All of the video sequences are subject to block-based random projection applied frame by frame; i.e., by partitioning each frame  $p$  into  $B \times B$  blocks and applying to each block an orthonormalized

dense Gaussian measurement matrix of size  $M_{B_p} \times B^2$  such that the subrate for frame  $p$  is  $S_p = M_{B_p}/B^2$ . Unless otherwise stated, we use a block size of  $B = 16$  and a GOP size of  $P = 8$  frames.

MC-BCS-SPL employs BCS-SPL for reconstruction of individual MC residual frames as well as in the multihypothesis initialization of Figure 3.3. For this BCS-SPL reconstruction, we use a 2D DCT as the transform operator and hard thresholding, for simplicity. For the ME/MC process in MC-BCS-SPL, we use full-search ME with quarter-pixel accuracy and a search window of  $\pm 15$  pixels. MAX\_ITERATIONS in Figure 3.2 is set to 5.

We compare MC-BCS-SPL to the intraframe-BCS-SPL and 3D-BCS-SPL alternatives. The most straightforward reconstruction is intraframe BCS-SPL (we use a block-based  $B \times B$  2D DCT as the transform basis). Additionally, we consider 3D-BCS-SPL with the 3D transform operator being a  $B \times B \times P$  3D DCT. Additionally, we note that 3D-BCS-SPL uses a GOP size of  $P = 4$  frames which we have observed empirically to yield results superior to a GOP size of  $P = 8$ .

As a primary measure of reconstruction quality, we calculate the peak signal-to-noise ratio (PSNR) averaged over all the frames under consideration. For MC-BCS-SPL, various subrates are employed for the key frames as well as the non-key frames; thus, we have two subrates in use: the key-frame subrate ( $S_K$ ), and the non-key-frame subrate ( $S_{NK}$ ). First, we consider the case wherein all frames have the same subrate, i.e.,  $S_K = S_{NK}$ . Alternatively, we also consider the case wherein the key frames have an increased subrate with respect to the non-key frames, i.e.,  $S_K > S_{NK}$ . A summary of the results from both cases is presented in Tables 3.2.

Table 3.2

Average PSNR in dB for several video sequences;  $S_K = 0.7$ , 88 frames

<i>Algorithm</i>	$S_{NK}$				
	0.1	0.2	0.3	0.4	0.5
Coastguard					
MC-BCS-SPL ( $S_K > S_{NK}$ )	<b>24.2</b>	<b>27.1</b>	<b>29.6</b>	<b>32.2</b>	<b>34.7</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	23.1	25.8	28.3	30.9	33.6
3D-BCS-SPL	22.8	24.8	26.4	27.9	29.6
Intraframe MS-BCS-SPL	22.7	24.7	26.4	28.0	29.6
Football					
MC-BCS-SPL ( $S_K > S_{NK}$ )	20.7	<b>23.4</b>	<b>25.9</b>	<b>28.1</b>	<b>30.1</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	<b>20.9</b>	<b>23.4</b>	25.5	27.4	29.4
3D-BCS-SPL	20.8	22.7	24.4	25.9	27.5
Intraframe MS-BCS-SPL	20.7	22.4	24.1	25.7	27.3
Foreman					
MC-BCS-SPL ( $S_K > S_{NK}$ )	<b>28.0</b>	<b>32.3</b>	<b>35.0</b>	<b>37.4</b>	<b>39.8</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	27.6	31.0	33.5	35.8	38.0
3D-BCS-SPL	26.5	29.2	31.4	33.3	35.1
Intraframe MS-BCS-SPL	26.0	28.8	31.2	33.2	35.1
Hall Monitor					
MC-BCS-SPL ( $S_K > S_{NK}$ )	<b>26.4</b>	<b>30.8</b>	<b>33.2</b>	<b>35.5</b>	<b>38.1</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	22.8	25.3	27.4	29.4	31.3
3D-BCS-SPL	22.8	25.2	27.3	29.1	31.0
Intraframe MS-BCS-SPL	22.6	24.8	26.9	28.8	30.7
Mother and Daughter					
MC-BCS-SPL ( $S_K > S_{NK}$ )	<b>33.9</b>	<b>38.3</b>	<b>40.7</b>	<b>42.8</b>	<b>45.0</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	30.7	34.0	36.6	38.9	41.0
3D-BCS-SPL	30.4	33.3	35.6	37.5	39.3
Intraframe MS-BCS-SPL	30.1	33.1	35.6	37.5	39.3
Susie					
MC-BCS-SPL ( $S_K > S_{NK}$ )	30.6	<b>34.8</b>	<b>37.3</b>	<b>39.5</b>	<b>41.5</b>
MC-BCS-SPL ( $S_K = S_{NK}$ )	<b>31.2</b>	34.4	36.8	38.8	40.8
3D-BCS-SPL	30.4	32.9	34.8	36.4	38.0
Intraframe MS-BCS-SPL	30.0	32.8	34.9	36.5	38.0

### 3.4.1 MC-BCS-SPL with Equal Subrate

The simplest situation is when all frames of a GOP, including the key frame, have an identical subrate. In this case, the average subrate per frame is  $S = S_K = S_{NK}$ . The PSNR performance of MC-BCS-SPL is measured at various subrates and compared with intraframe BCS-SPL and 3D-BCS-SPL in Figures 3.6 and 3.7. In these graphs, we see that 3D-BCS-SPL yields similar or slightly superior performance as compared to intraframe BCS-SPL; this is as expected because the 3D reconstruction exploits what frame-to-frame correlation exists in the sequence. However, the advantage is not very large in most instances, so it would appear insufficient to reconstruct based simply on frame-to-frame correlation.

On the other hand, ME/MC in MC-BCS-SPL can largely track object motion and, hence, can increase compressibility in the sense of (2.4) in the resulting residual to yield higher-quality reconstruction. It is notable that MC-BCS-SPL achieves a 3–5 dB gain at the higher subrates for some sequences (“Coastguard,” “Foreman,” “Mobile,” “Stefan”), as can be observed in Table 3.2.

### 3.4.2 MC-BCS-SPL with Key Frames of Increased Subrate

Since the key frames constitute only a small number of the total frames in a sequence, and they serve somewhat as “anchors” to the forward/backward ME process of reconstruction in a GOP, it is reasonable to consider the situation in which key frames are given a higher subrate than the non-key frames such that they are reconstructed with high quality. Thus, we perform a battery of experiments which measure the PSNR when the subrate

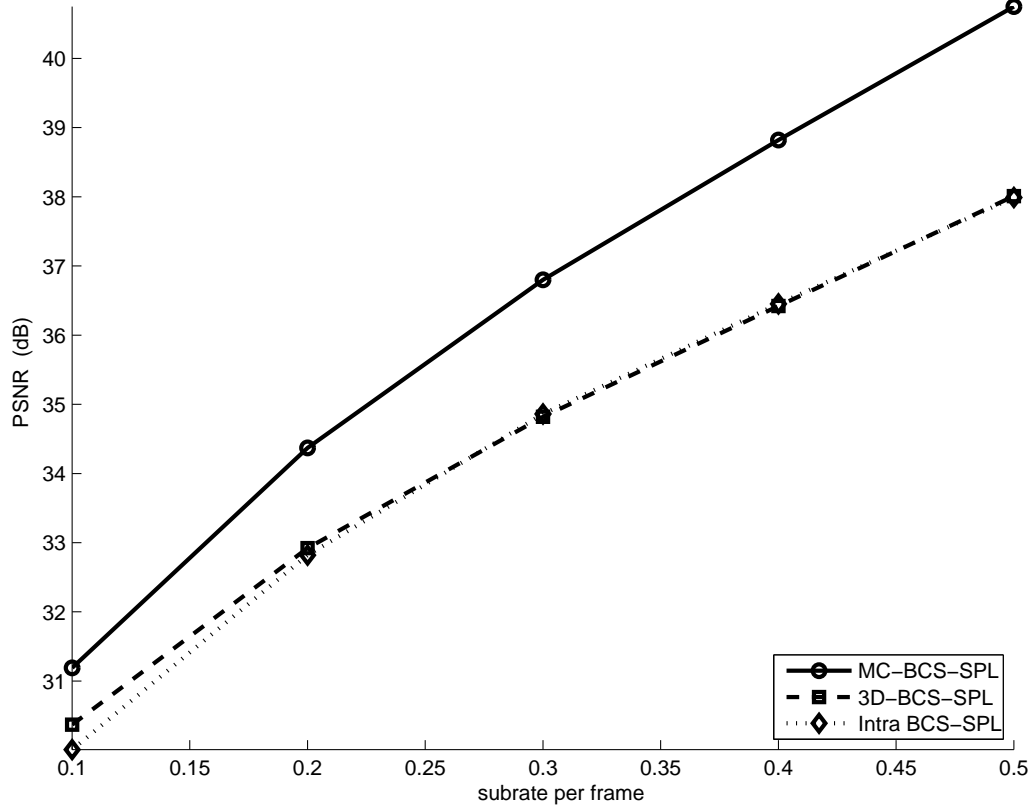


Figure 3.6

Performance of MC-BCS-SPL on the “Susie” sequence for equal subrate,  $S_K = S_{NK}$ .  
PSNR is averaged over all frames of the sequence.

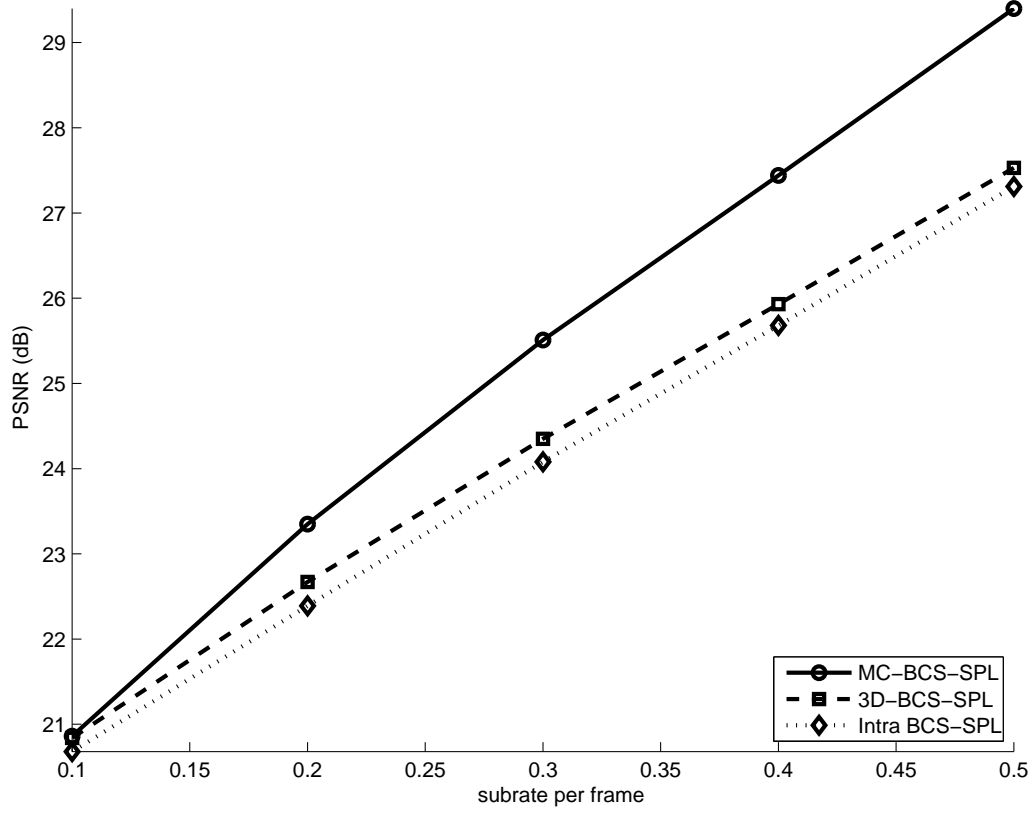


Figure 3.7

Performance of MC-BCS-SPL on the “Football” sequence for equal subrate,  $S_K = S_{NK}$ .  
PSNR is averaged over all frames of the sequence.

for key frames is increased beyond that of the non-key frames. In this situation, we set a target average subrate per frame to be  $S$ —we will have  $S$  range from 0.1 to 0.5 in the experiments. We then set the key-frame subrate to be incrementally higher than  $S$ ; i.e., we use

$$S_K = S + 0.4, \quad (3.5)$$

which we have found works well in practice. Then, the corresponding non-key subrate needed to produce the desired target  $S$  is

$$S_{NK} = \frac{S \cdot P - S_K}{P - 1}, \quad (3.6)$$

where  $P$  is the GOP size.

Figures 3.8 and 3.9 depict the performance of MC-BCS-SPL with increased  $S_K$ , where “MC-BCS-SPL (unequal)” refers to MC-BCS-SPL with  $S_K > S_{NK}$ ; “MC-BCS-SPL (equal)” refers to MC-BCS-SPL with  $S_K = S_{NK}$ . In these graphs, MC-BCS-SPL with increased  $S_K$  is compared to MC-BCS-SPL with  $S_K = S_{NK}$ , i.e., the “equal subrate” paradigm as well as to 3D-BCS-SPL (intraframe BCS-SPL is omitted due to similar performance to 3D-BCS-SPL). As expected, the higher-quality reconstruction of the key frames results in a significant performance improvement for MC-BCS-SPL. In particular, remarkable gains are exhibited for the “Hall Monitor” and “Mother and Daughter” sequences which did not show a large gain over intraframe BCS-SPL in the previous, equal-subrate experiments.

Additionally, visual quality of the intraframe BCS-SPL and MC-BCS-SPL recoveries of the “Football” and “Susie” sequences are compared in Figures 3.10 and 3.11 for the center frame ( $p = 4$ ) of the first GOP. As can be seen, MC-BCS-SPL works especially well on

static parts in the image due to ME/MC successfully reducing the residual energy between consecutive frames in such static regions. For more dynamic sequences, although MC-BCS-SPL still results in some degree of “mosquito noise” on moving objects, it provides visual quality significantly superior to that of intraframe BCS-SPL.



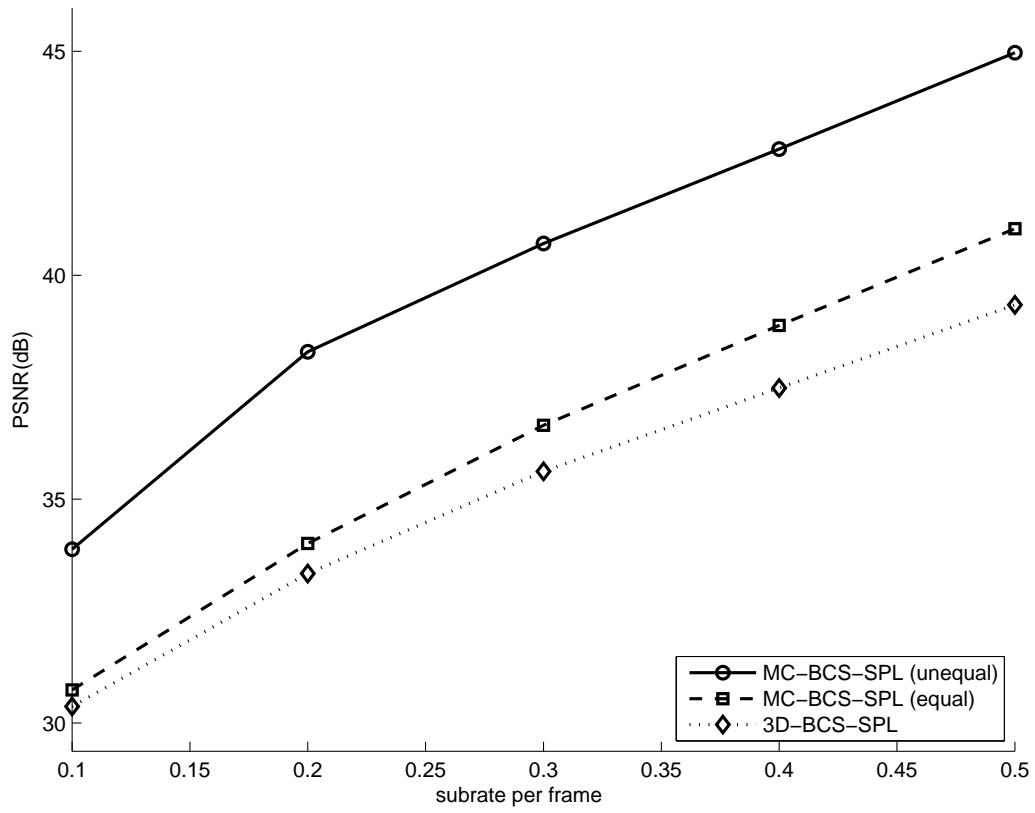


Figure 3.8

Performance of MC-BCS-SPL on the “Mother and Daughter” sequence for unequal subrate; PSNR is averaged over all frames of the sequence.

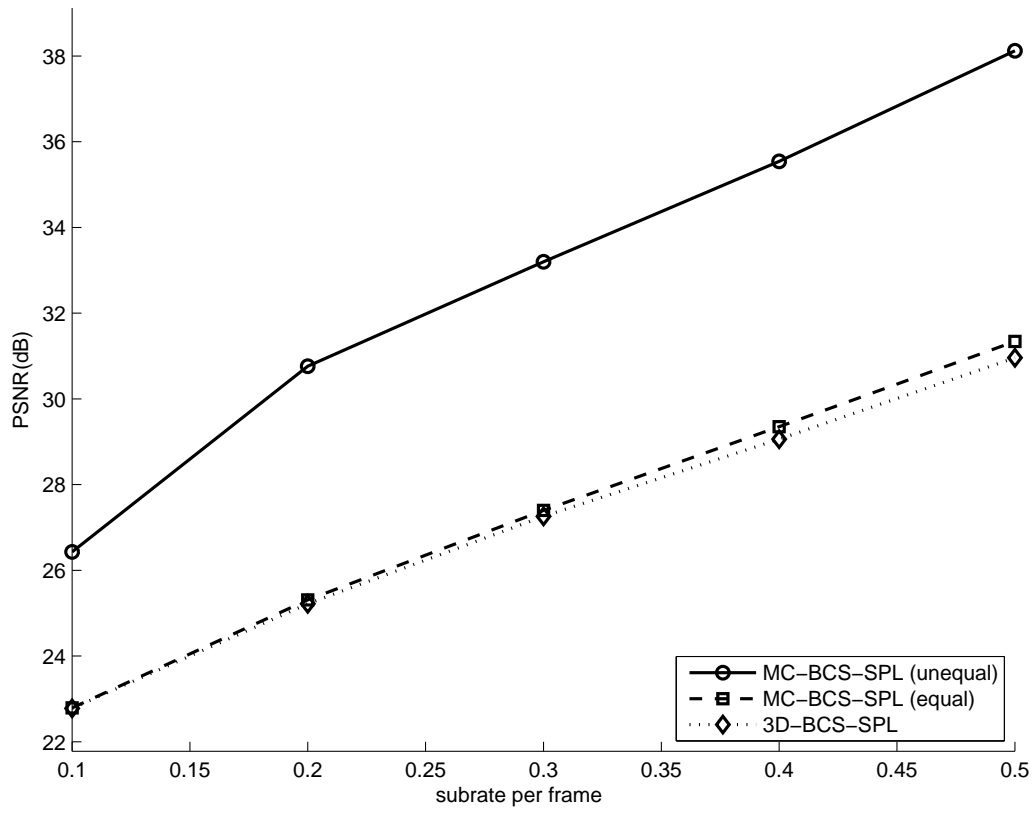


Figure 3.9

Performance of MC-BCS-SPL on the “Hall Monitor” sequence for unequal subrate;  
PSNR is averaged over all frames of the sequence.



(a)



(b)

Figure 3.10

Reconstructed center frame for “Football.” (a) Intraframe BCS-SPL with  $S = 0.3$ , 24.02 dB; (b) MC-BCS-SPL with unequal subrate and  $S = 0.3$  ( $S_K = 0.7$ , and  $S_{NK} = 0.24$ ).



(a)



(b)

Figure 3.11

Reconstructed center frame for “Susie.” (a) Intraframe BCS-SPL with  $S = 0.3$ , 32.48 dB; (b) MC-BCS-SPL with unequal subrate and  $S = 0.3$  ( $S_K = 0.7$ , and  $S_{NK} = 0.24$ ).

### 3.5 Comparison of Various CS Techniques for Video

We now present a comprehensive comparison between several CS reconstruction algorithms for video. We use the first 88 frames of the “Foreman,” “Coastguard,” “Hall Monitor,” and “Mother and Daughter” sequences. In all cases, we use a GOP size of  $P = 8$  frames with key frames starting each GOP having a subrate of  $S_K = 0.7$ . The intervening non-key frames have subrate  $S_{NK}$  varying between 0.1 and 0.5. It should be noted that, unlike in the presented in Section 3.4.2, fixed subrates for key frames and non-key frames are used for simplicity while maintaining the same experimental conditions for all approaches.

We compare MC-BCS-SPL to other techniques. Being block based, MC-BCS-SPL features block-based measurement in the spatial domain applied identically to each video frame; the block size is  $16 \times 16$ . An orthonormalized dense Gaussian measurement matrix is used for measurement, and a dual-tree discrete wavelet transform (DDWT) [73] is used as the sparsity transform.

We also consider three prominent CS reconstruction algorithms, Modified-CS-Residual [104], focal underdetermined system solver in k-t space (k-t FOCUSS) [70, 71], and distributed compressed video sensing (DISCOS) [33] (see [49] for a thorough overview of each). Briefly, k-t FOCUSS applies ME/MC of non-key frames based on the neighboring key frames in an iterative fashion; Modified-CS-Residual tracks the sparsity pattern over time without the use of ME/MC; and, finally, DISCOS performs non-iterative multihypothesis prediction of the current block within a spatial window using an  $\ell_1$  minimization.

We use the implementations of k-t FOCUSS<sup>1</sup>, Modified-CS-Residual<sup>2</sup>, and DISCOS<sup>3</sup> available from their respective authors. We note that both k-t FOCUSS and Modified-CS-Residual were originally intended for magnetic resonance imaging (MRI) applications; as such, they are both employ frame-by-frame measurement driven by a 2D Fourier transform applied to each frame with higher subrate devoted to low-frequency coefficients. Although there are a number of other reconstruction algorithms for video in the literature, none of these, to our knowledge, have implementations readily available at the time of this writing. As a consequence, we present results only for those algorithms identified above.

Table 3.3 tabulates the observed results with visual results for the 4th frame of the GOP of the “Foreman” and “Mother-and-Daughter” sequences given in Figure 3.12 and Figure 3.12, respectively. As can be seen notably in Table 3.3, the iterative ME/MC-driven MC-BCS-SPL usually outperforms the other techniques considered, sometimes by as much as 2–3 db.

Although none of the implementations have been particularly optimized for execution speed, we present reconstruction times for the algorithms in Table 3.4. Here, we measure the typical length of time required to reconstruction one frame out of the sequence. The results shows MC-BCS-SPL runs about three to four times longer than DISCOS or k-t FOCUSS; however, MC-BCS-SPL is significantly faster than Modified-CS-Residual.

---

<sup>1</sup>[http://bisp.kaist.ac.kr/research\\_02.htm](http://bisp.kaist.ac.kr/research_02.htm)

<sup>2</sup><http://home.engineering.iastate.edu/~luwei/modcs/>

<sup>3</sup>Available directly from T. Do.

Table 3.3

Average PSNR in dB for several video sequences;  $S_K = 0.7$ , 88 frames

<i>Algorithm</i>	$S_{NK}$				
	0.1	0.2	0.3	0.4	0.5
Coastguard					
MC-BCS-SPL	25.8	<b>28.1</b>	<b>29.8</b>	<b>31.3</b>	<b>33.0</b>
k-t FOCUSS	24.3	26.2	27.8	29.3	29.2
Modified-CS-Residual	23.7	25.3	26.8	28.3	30.0
DISCOS	<b>26.2</b>	27.4	28.7	29.6	30.7
Football					
MC-BCS-SPL	<b>21.9</b>	<b>24.9</b>	<b>27.0</b>	<b>28.6</b>	<b>30.2</b>
k-t FOCUSS	20.7	22.8	24.8	27.3	27.1
Modified-CS-Residual	20.8	22.8	24.3	26	27.6
DISCOS	21.5	22.9	24.1	25.4	26.4
Foreman					
MC-BCS-SPL	31.2	<b>34.1</b>	<b>36.2</b>	<b>38.1</b>	<b>39.7</b>
k-t FOCUSS	27.2	29.4	32.4	34.1	33.9
Modified-CS-Residual	25.7	27.4	29.0	30.6	32.3
DISCOS	<b>31.4</b>	32.7	33.9	34.9	35.6
Hall Monitor					
MC-BCS-SPL	31.8	<b>34.8</b>	<b>35.8</b>	36.4	<b>36.9</b>
k-t FOCUSS	<b>32.9</b>	34.3	35.7	<b>36.6</b>	36.6
Modified-CS-Residual	26.4	27.6	28.9	30.2	31.7
DISCOS	32.4	33.6	34.7	35.8	36.4
Mother and Daughter					
MC-BCS-SPL	<b>38.3</b>	<b>40.5</b>	<b>41.7</b>	<b>42.6</b>	<b>43.4</b>
k-t FOCUSS	36.1	37.5	39.1	40.3	40.2
Modified-CS-Residual	30.0	31.4	32.8	34.3	35.9
DISCOS	36.7	37.8	39.0	39.9	40.6
Susie					
MC-BCS-SPL	33.2	<b>36.8</b>	<b>38.8</b>	<b>40.3</b>	<b>41.6</b>
k-t FOCUSS	31.0	33.9	35.4	37.6	37.9
Modified-CS-Residual	29.1	31.1	32.8	34.6	36.3
DISCOS	<b>33.8</b>	35.1	36.1	37.2	27.8



Figure 3.12

Reconstructions of frame 4 of the “Foreman” sequence for  $S_K = 0.7$  and  $S_{NK} = 0.3$ .





Figure 3.13

Reconstructions of frame 4 of the “Mother and Daughter” sequence for  $S_K = 0.7$  and  $S_{NK} = 0.3$ .

Table 3.4

Reconstruction time in seconds per frame (spf)

<i>Algorithm</i>	<i>Time (spf)</i>
DISCOS	41
k-t FOCUSS	46
MC-BCS-SPL	159
Modified-CS-Residual	699

### 3.6 Remarks

Because of its multi-dimensional nature, the size of the CS problem grows quickly for video. Furthermore, it is less than clear whether global CS acquisition in both time and spatial can be implemented in practice. As a consequence, frame-by-frame acquisition is typically employed for video, and we followed this strategy throughout this chapter. Furthermore, capitalizing upon the BCS-SPL architecture of Chapter 2 provides fast and simple yet high-quality image reconstruction as a foundation for video recovery.

The main contribution of this chapter is the development of a motion-compensated version of BCS-SPL for use with the CS reconstruction of video. In this MC-BCS-SPL, ME/MC predictions were used to create a residual of the frame to be recovered. Such a residual is typically more compressible—in the sense of (2.4)—than the original frame, rendering it more amenable to CS recovery. Furthermore, by using a higher subrate for key frames such that they carry more information and thus “anchor” the reconstruction of the remainder of the sequence, we obtained more accurate motion vectors, resulting in higher-quality predicted images. As a consequence, considerable gain was achieved as compared to using an identical subrate for all frames. We demonstrated that MC-BCS-

SPL significantly improved distortion performance over a straightforward recovery of the frames independently, a 3D joint reconstruction, and several alternative CS reconstructions for video from prior literature, including DISCOS [33], k-t FOCUSS [70, 71], and Modified-CS-Residual [104].

In the next chapter, we return to the issue of the CS of images, considering the addition of quantization to the CS measurement process in order to produce a true bitstream from the CS measurements. We note that, while our focus will be on images exclusively, it would be possible to extend the methodology to the MC-BCS-SPL video reconstruction that we proposed in this chapter.

## CHAPTER 4

### DPCM FOR QUANTIZED BLOCK-BASED COMPRESSED SENSING OF IMAGES

In the field of information theory as established by Shannon [91, 93], *source coding* can be considered to be “the conversion of arbitrary signals into an efficient digital representation” [52]; this digital representation is often called a *bitstream*. Since, in most cases, it is desired that the resulting bitstream use as few bits as possible, source coding has often been referred to as *data compression*. The performance of such data compression is measured in terms of a *bitrate*—i.e., the number of bits contained in the bitstream per original source sample. The compression may result in an exact representation of the original source, in which case the compression is *lossless*; otherwise, it is *lossy*. In the case of lossy compression, performance is also measured in terms of the quality of the representation using some fidelity measure, typically called the *distortion*. Thus, in information theory, the rate-distortion performance characterizes the compression in the form of a tradeoff between bitrate and distortion.

However, in the context of compressed sensing (CS), we have considered thus far only subrate-distortion performance. This may be considered in certain respects to be unrealistic, since the dimensionality reduction obtained by the CS measurement process does not produce compression in the strict information-theoretic sense; indeed, some form of quantization is necessary to produce a compressed bitstream from the CS measurements.

Although such quantization is inevitable in any real-life implementation of a CS measurement scheme, CS literature has largely avoided the topic of quantization.

In this chapter, we study the recovery of the quantized CS measurements. We aim to find a method for quantized CS without placing an undue burden on the sensor device in order to maintain a reduced sensor complexity, which is the central premise of CS. To this end, we once again invoke block-based compressed sensing (BCS).

Specifically, we saw in Section 2.5 that the measurement matrix of BCS takes a block-diagonal form; if the same block measurement matrix is used for every block, the resulting measurements possess a high degree of correlation from block to block. In order to exploit such correlation between the block measurements, we adopt differential pulse-code modulation (DPCM) (widely used in analog-to-digital conversion and video coding) as a form of predictive coding. The proposed algorithm combines DPCM with uniform scalar quantization (SQ) to provide a simple and effective block-based quantized CS of images with only a small complexity cost to the simple CS sampler. Experimental results demonstrate significant improvement in rate-distortion performance as compared SQ used alone in several block-based CS reconstruction algorithms. Additionally, rate-distortion performance superior to that of alternative quantized-CS techniques relying on optimized quantization or reconstruction is observed. We note that the work that is presented in this chapter was previously published as [84].

## 4.1 Background

In this section, some of the techniques for quantized CS measurements are considered and the motivation is explained. Also, DPCM, the algorithm used in the quantized BCS measurements, in general image coding framework is reviewed and the reason why DPCM works in the BCS framework is explained.

### 4.1.1 Quantized Compressed Sensing

The straightforward solution to incorporating quantization into the CS paradigm is simply to apply SQ to each of the CS measurements produced by the sensing device. However, it has been established that such an SQ-based solution is highly inefficient in terms of information-theoretic rate-distortion performance as compared to traditional source-coding techniques (e.g., [54]). As a consequence, there have been a variety of efforts in recent literature aimed at the improvement of rate-distortion performance of quantized CS, largely through an optimization of the quantization process (e.g., [27, 96]), the reconstruction process (e.g., [16, 65]), or both (e.g., [108, 109]).

In [27] and [96], non-uniform quantization optimized with respect to mean square error (MSE) demonstrated rate-distortion performance superior to that of simple SQ, but finding an optimal quantizer required an iterative process which significantly increased computation complexity within the CS encoding procedure. Alternatively, the basis pursuit dequantizer (BPDQ) of [65] found a sparse solution from scalar-quantized measurements by adding particular data-fidelity constraints to enhance reconstruction quality. Similarly, in [16], so-called “1-bit CS” recovered the signal using only sign information of the measure-

ments, again through addition of a sign-consistency constraint. Both BPDQ and 1-bit CS work well in the case where the number of measurements are sufficiently large to capture sparsity in the signal ( $M/K > 10$ ); however, images are sometimes quite complex (i.e., the  $K$ -sparsity is high), so they are not well suited for such reconstruction. Finally, [108, 109] used a method called progressive quantization (PQ) that separates measurements into two different layers (base and refinement) and performs coarse and fine quantization, respectively, at the encoder side, with the initial reconstruction of the base layer being used to determine some of the least-significant bits of the measurements in the refinement layer under the assumption that measurements from the same random matrix have the same probabilistic distribution. PQ works well with a certain class of images, but, as shown Section 4.3.3, the approach does not work well on complex image textures (i.e., the Barbara image).

In this chapter, in contrast to such prior work on quantized CS which largely relies on optimized quantization or reconstruction, we propose a straightforward process of quantization via simple uniform SQ applied in conjunction with DPCM of the CS measurements. Our framework is applicable only to the CS of images effectuated in blocks, i.e., BCS [24, 48–50, 82]. In essence, at the sensor side of the system, rather than applying quantization directly to each block of CS measurements, a prediction of the block is made and subtracted from the current block of measurements in the measurement domain. The resulting residual is then scalar-quantized. At the reconstruction side of the system, the same prediction is added onto the dequantized residuals to produce the quantized CS measurements ready for BCS-based reconstruction.

We mention that our proposed DPCM-based approach bears some resemblance to the sigma-delta quantization for CS in [55] in the sense that sigma-delta modulation also provides quantization based on differences between CS measurements. Our approach is somewhat simpler conceptually ([55] requires Sobolev frames for reconstruction); additionally, we apply DPCM across blocks rather than sample-by-sample as in the sigma-delta modulation of [55]. Another difference is that we study the effect on rate-distortion performance while sigma-delta method concentrates on alleviating quantization error regardless rate.

#### 4.1.2 DPCM for Natural Images

Generally, images possess spatial redundancies. If there exists high correlation between pixels in an image, or if one pixel is predictable based on neighboring pixels, then such redundancies can be exploited by predictive coding. One method is DPCM which transmits pixel differences rather than the pixel values themselves, thereby increasing rate-distortion efficiency due to the fact that pixel differences are near zero with high probability and thus require fewer bits to code.

Specifically, consider a discrete sequence,  $\mathbf{x} = \begin{bmatrix} x[1] & \dots & x[n] & \dots & x[N] \end{bmatrix}^T$ , where  $N$  is the total number of elements, or pixels, of a rasterized image. An estimation of current pixel,  $\tilde{x}[n]$ , can be obtained by linear prediction of the previous pixels, which can be expressed as

$$\tilde{x}[n] = \sum_{i=1}^k a_i x[n-i], \quad (4.1)$$

where  $k$  is the order of prediction coefficients. For example, if  $k = 1$  and  $a_1 = 1.0$ , then the estimate of the current pixel becomes simply the previous pixel value,  $\tilde{x}[n] = x[n-1]$ .



The pixel difference  $d[n]$  is the difference between the current pixel,  $x[n]$ , and the estimate of the current pixel,  $\tilde{x}[n]$ ,

$$d[n] = x[n] - \tilde{x}[n]. \quad (4.2)$$

In DPCM, the pixel difference is quantized to produce the index for transmission,

$$i[n] = Q[d[n]], \quad (4.3)$$

where  $Q[\cdot]$  can be any type of quantizer, i.e., a uniform scalar quantizer or an MSE-based optimized non-uniform quantizer. At the decoder, quantized pixel-difference value,  $\hat{d}[n]$ , is recovered from the transmitted indices, and the current pixel is reconstructed by adding the predicted current pixel,  $\tilde{x}$ , to the pixel difference,

$$\hat{x}[n] = \hat{d}[n] + \tilde{x}[n]. \quad (4.4)$$

It should be noted that, because the indices for the quantized difference values are transmitted, the reconstructed pixel,  $\hat{x}[n]$ , relies on the previously reconstructed pixels,  $\hat{x}[n-1], \dots, \hat{x}[n-k]$ . Therefore, the same feedback loop exists in both the encoder and the decoder in order to generate the predicted pixel based on the reconstructed values of the quantized pixel differences. The overall process is depicted in Figure 4.1.

The reconstructed pixel error  $e[n]$  is the difference between current input pixel and its corresponding reconstruction,

$$e[n] = x[n] - \hat{x}[n]. \quad (4.5)$$

From (4.2), (4.4), and (4.5), we have

$$\begin{aligned} e[n] &= x[n] - (\hat{d}[n] + \tilde{x}[n]) = x[n] - (\hat{d}[n] + x[n] - d[n]) \\ &= d[n] - \hat{d}[n]. \end{aligned} \quad (4.6)$$

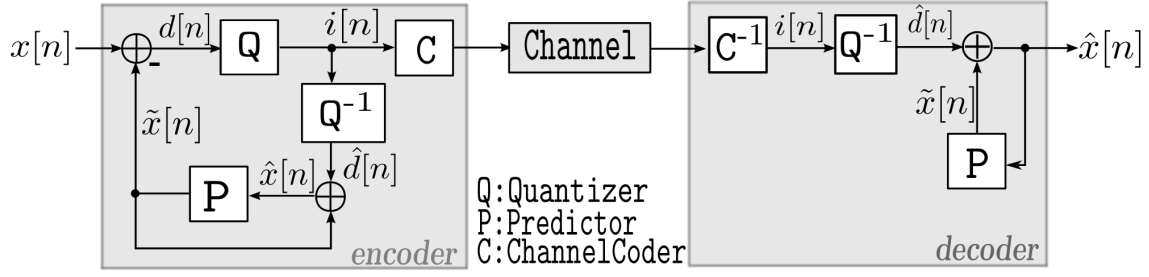


Figure 4.1

DPCM for an image.  $C$  is any entropy coder, such as arithmetic coding.

We see that (4.6) implies that the reconstruction error is the same as the quantization error, which means that DPCM does not produce any additional error beyond the quantization error. Therefore, given a quantizer, the purpose of using DPCM is to reduce the bit rate.

Finally, in order to find an optimal predictor, we choose  $a_i$  in (4.1) to minimize the MSE between the input and predicted pixels over the whole image. Assuming zero mean, we have

$$\begin{aligned}\sigma_d^2 &= E[d[n]^2] = E[(x[n] - \tilde{x}[n])^2] \\ &= E\left[\left(x[n] - \sum_{i=1}^k a_i x[n-i]\right)^2\right],\end{aligned}\tag{4.7}$$

where  $E[\cdot]$  is expectation. After taking partial derivatives of the cost function with respect to each  $a_i$  and setting to zero, we arrive at

$$\begin{aligned}\frac{\partial \sigma_d^2}{\partial a_i} &= -2E\left[\left(x[n] - \sum_{i=1}^k a_i x[n-i]\right)x[n-i]\right], 1 \leq i \leq k \\ &= 0.\end{aligned}\tag{4.8}$$

Thus, the optimal prediction coefficients are found from (4.8) if the  $E[x[n]x[n-i]]$  (the autocovariances) are known. Using an orthogonality condition (the optimal linear predic-

tion is orthogonal to the current input,  $E[(x[n] - \tilde{x}[n])\tilde{x}[n]] = 0$ ), the minimum prediction error using the optimal predictor can be found to be (see, e.g., [85])

$$(\sigma_d^2)_{min} = E[(x[n] - \tilde{x}[n])^2] = E[(x[n] - \tilde{x}[n])x[n]]. \quad (4.9)$$

For the first-order predictor, for example, the optimal value is

$$a_1 = \frac{E[x[n]x[n-1]]}{E[(x[n] - 1)^2]} = \frac{R_x(1)}{R_x(0)} = \rho_1, \quad (4.10)$$

where  $R_x(0)$  and  $R_x(1)$  are the autocorrelation function of the input sequence at time lag 0 and 1, respectively, and  $\rho_1$  is the correlation coefficient. The minimum MSE for the first-order linear predictor is

$$\begin{aligned} (\sigma_d^2)_{min} &= E[(x[n] - \tilde{x}[n])^2] = E[(x[n] - ax[n-1])x[n]] \\ &= R_x(0) - aR_x(1) = R_x(0)(1 - \rho_1^2). \end{aligned} \quad (4.11)$$

With the zero-mean assumption, (4.11) becomes

$$\sigma_d^2 = \sigma_x^2(1 - \rho_1^2). \quad (4.12)$$

The consequence of this derivation is that the variance of the difference signal,  $d[n]$ , is reduced from that of signal itself,  $x[n]$ , by the factor of  $(1 - \rho_1^2)$  when we use the optimal first-order linear predictor. This reduction in variance can be represented as a prediction gain,  $G_p$ , defined as the ratio of the input signal variance to the differential signal variance (see, e.g., [68]),

$$G_p = \frac{\sigma_x^2}{\sigma_d^2} = \frac{1}{(1 - \rho_1^2)}. \quad (4.13)$$

Under an assumption of Gaussianity, the corresponding reduction in bit rate can be shown to be

$$\Delta R = R_{\text{SQ}} - R_{\text{DPCM}} = \frac{1}{2} \log_2 \frac{\sigma_x^2}{\sigma_e^2} - \frac{1}{2} \log_2 \frac{\sigma_d^2}{\sigma_e^2} = \frac{1}{2} \log_2 \frac{\sigma_x^2}{\sigma_d^2}, \quad (4.14)$$

where  $\sigma_e^2$  is the variance of the quantization error. Again, using first-order linear prediction, we have

$$\Delta R = \frac{1}{2} \log_2 \frac{1}{(1 - \rho_1^2)}. \quad (4.15)$$

Typically, grayscale images have  $\rho_1 \approx 0.95$ ; assuming a large number of quantization levels, the reduction can be around 1 to 3 bits per pixel (bpp) [67]. The prediction gain specific to BCS will be examined in Section 4.2.2.

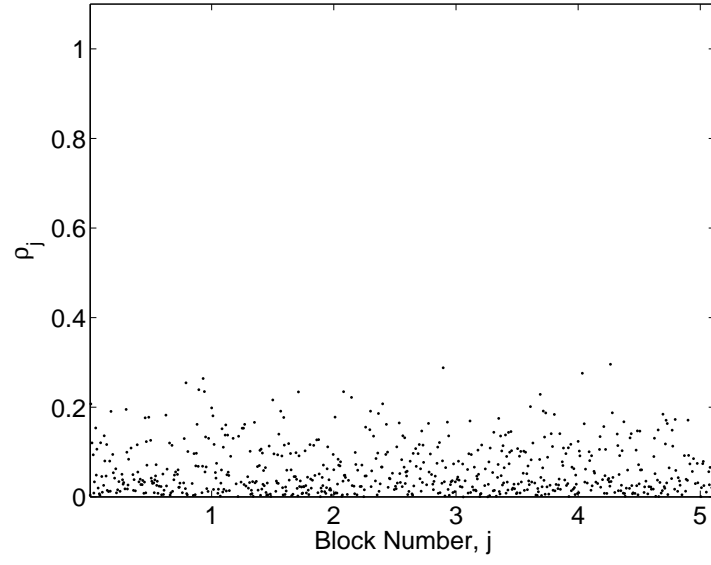
### 4.1.3 Correlation of Measurements in BCS

In general, DPCM works when signals possess a significant degree of correlation from one time to the next. Such correlation typically exists in images from one image block to the next and in video from one frame to another temporally. However, such correlation does not exist in CS because the random projection process renders measurement-to-measurement correlation very low. An example is shown in Figure 4.2(a). A rasterized  $512 \times 512$  grayscale image is measured using a single (non-diagonal) measurement matrix,  $\mathbf{y} = \Phi \mathbf{x}$ , with a subrate of 0.5, and then the vector of measurements,  $\mathbf{y}$ , is partitioned into 512 blocks of size  $1 \times 256$ . We then calculate the block-to-block correlation by defining the measurement-domain correlation coefficient between blocks  $\mathbf{y}^{(j)}$  and  $\mathbf{y}^{(j-1)}$  as

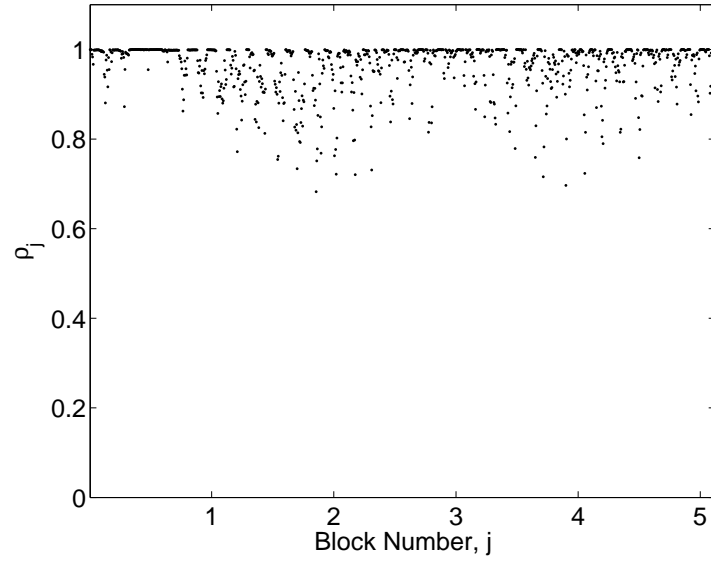
$$\rho_j = \frac{\mathbf{y}^{(j)T} \mathbf{y}^{(j-1)}}{\|\mathbf{y}^{(j)}\| \|\mathbf{y}^{(j-1)}\|}. \quad (4.16)$$

As illustrated in Figure 4.2(a), the normalized correlation coefficients of CS measurements are mostly scattered in the range 0 to 0.2. The average correlation over all groups is  $\bar{\rho} = 0.026$ . This implies that the measurement process of CS reduces the autocorrelation of the pixels (note that raw image pixels typically have a correlation of  $\rho \approx 0.95$  [67]), and, therefore, it is difficult to use DPCM directly in such a framework because of the low correlation.

However, if the random-projection process is BCS such that  $\mathbf{y}^{(j)} = \Phi_B \mathbf{x}^{(j)}$ , where  $j$  is the index of the image block, then the measurements preserve correlation because the measurement matrix  $\Phi$  is composed of multiple  $\Phi_B$  matrices which are the same. The correlation coefficients of BCS measurements are also given by (4.16). Figure 4.2(b) shows the normalized correlation coefficients of the measurements of the same image above (also for a subrate of 0.5); we see that, for many blocks,  $\rho_j$  is close to 1.0, while the average correlation over all blocks,  $\bar{\rho} = 0.971$ , indicates that the consecutive blocks of measurements are typically highly correlated, even if the measurements themselves are not. Based on this correlation information in BCS, one can hope to achieve the DPCM rate reduction anticipated by (4.15). In the next section, we consider the specifics of applying DPCM in the BCS framework.



(a) CS,  $\bar{\rho} = 0.026$



(b) BCS,  $\bar{\rho} = 0.971$

Figure 4.2

Correlation coefficients of measurements in CS and BCS for the  $512 \times 512$  Lenna image.

## 4.2 DPCM for Quantized BCS-SPL

Among the many existing image-coding techniques, DPCM has the special advantage of a relatively simple implementation. This feature is well-suited to the CS paradigm which features, as a primary goal, an inexpensive signal-measurement process. There are two design considerations for DPCM for CS—the quantizer and the predictor. In this dissertation, we only use a uniform scalar quantizer even though it is shown that MSE-based non-uniform quantizers can achieve lower distortion [27] because the required optimization process is too burdensome for the encoder. Furthermore, the non-uniform quantizer itself (typically in the form of a table of quantization bins) needs to be sent to the decoder, entailing an additional, out-of-band rate cost. While there exists an alternative—table look-up using prior knowledge that pixel differences have a Laplace distribution which can be pre-calculated using an MSE-based non-uniform quantizer, such as a Lloyd-Max quantizer—we observed that the gain in rate-distortion performance is marginal for such an architecture.

As for the predictor, we considered two types of linear predictions, 1D linear prediction and 2D linear prediction. DPCM with 1D linear prediction makes use of the correlation of adjacent pixels within the same scan line. On the other hand, 2D linear prediction exploits line-to-line correlation. We denote these approaches 1D-DPCM and 2D-DPCM, respectively.

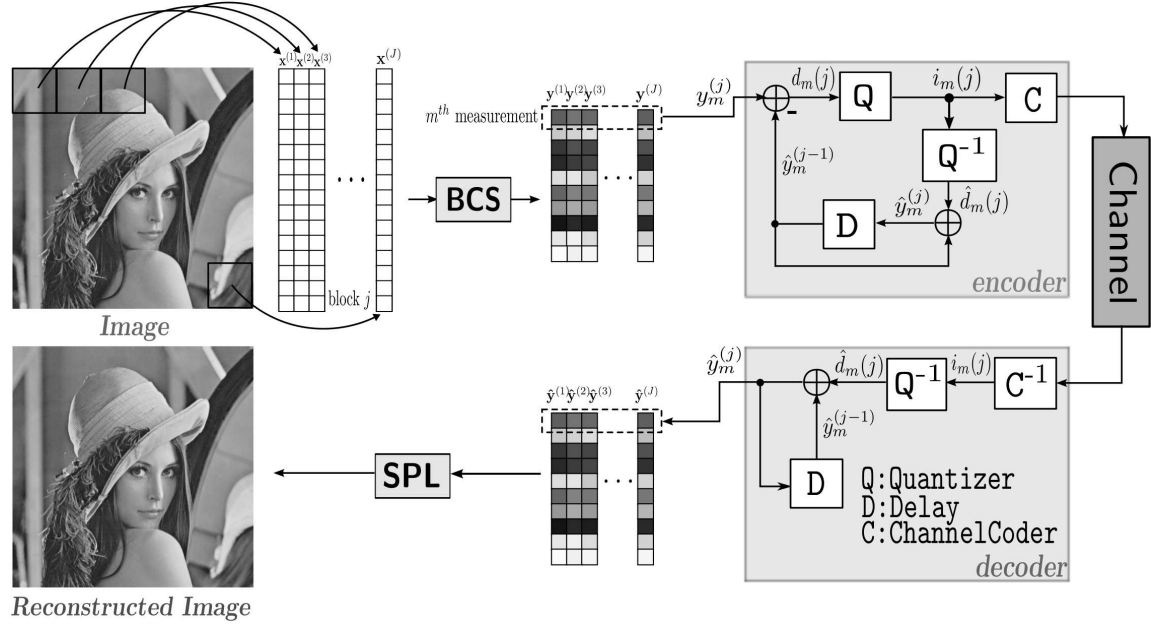


Figure 4.3

Application of DPCM and SQ to the BCS-SPL architecture. BCS is implemented with any CS-based image-acquisition, such as the single-pixel camera of [41].

#### 4.2.1 1D-DPCM for Quantized BCS-SPL

Effectively, our proposed approach for 1D-DPCM applies DPCM and SQ onto the block-based CS measurements within the block-based compressed sensing with smooth projected Landweber reconstruction (BCS-SPL) architecture of Section 2.5 as shown in Figure 4.3, where  $Q$  is uniform SQ,  $D$  is a single-block delay buffer, and  $C$  is any entropy coder, such as arithmetic coding.



On the sensor side of the system, BCS measurements are acquired as usual via (2.22) using  $B \times B$  blocks from the original image, producing  $M$ -dimensional measurement vector

$$\mathbf{y}^{(j)} = \begin{bmatrix} y_1^{(j)} & \cdots & y_m^{(j)} & \cdots & y_{M_B}^{(j)} \end{bmatrix}^T = \Phi_B \mathbf{x}^{(j)} \quad (4.17)$$

for block  $j$  of the image,  $\mathbf{x}^{(j)}$ . For component  $m$  in measurement vector  $\mathbf{y}^{(j)}$ , a prediction is subtracted and the residual is scalar-quantized. Specifically, to predict  $y_m^{(j)}$ , we use the corresponding vector component of the previously processed block  $\hat{\mathbf{y}}^{(j-1)}$ . That is, the residual

$$d_m^{(j)} = y_m^{(j)} - \hat{y}_m^{(j-1)} \quad (4.18)$$

is scalar-quantized to produce quantization index

$$i_m^{(j)} = Q \left[ d_m^{(j)} \right] \quad (4.19)$$

which is then entropy coded. The DPCM feedback loop consists of dequantization of  $i_m^{(j)}$ , producing the quantized residual,

$$\hat{d}_m^{(j)} = Q^{-1} \left[ i_m^{(j)} \right] \quad (4.20)$$

such that

$$\hat{y}_m^{(j)} = \hat{d}_m^{(j)} + \hat{y}_m^{(j-1)}. \quad (4.21)$$

We note that the set of measurements in the first block is processed in the same manner by initializing  $\hat{\mathbf{y}}^{(0)}$  to be the zero vector. Finally, we use a one-block delay for the predictor; for example,  $k = 1$  and  $a_1 = 1.0$  in (4.1). Even though it is possible to calculate the optimal prediction coefficients  $a_i$  by using correlation information of the measurements

such as in (4.10), this process requires additional computation. The use of an approximate correlation coefficient such as  $\rho \approx 0.95$  is also possible, but the difference in performance was observed to be minimal.

The basic approach of Figure 4.3, which illustrates our proposed framework for the incorporation of DPCM into BCS-SPL, can also be applied to multiscale BCS-SPL (MS-BCS-SPL) (Section 2.6) as well as the multihypothesis variant of BCS-SPL in [24]. Specifically, for the former, wavelet-domain blocks in the baseband are fed into the DPCM encoder, while the other subbands are quantized directly with uniform SQ (unlike the baseband coefficients, those in the other subbands have low correlation). For the latter, multihypothesis BCS-SPL (MH-BCS-SPL) consists of multiple predictions culled from the image being reconstructed, following which reconstruction is driven by the measurement-domain residual resulting from the predictions. To incorporate DPCM into MH-BCS-SPL, the DPCM encoder simply processes the same measurements as those of the original BCS-SPL (without reserving a holdset for cross-validation as was described in [24]).

Finally, the block size is chosen empirically. Without quantization, BCS-SPL typically yields better results with large block sizes. Considering the memory allocation of the measurement matrix, reasonable choices are block sizes of  $16 \times 16$  to  $32 \times 32$  as were used in previous chapters for BCS-SPL of video and images, respectively. However, quantized BCS-SPL is no longer measured by subrate-distortion performance but by rate-distortion performance. Therefore, it is necessary to reconsider parameters such as a block size to achieve good rate-distortion performance.

Figure 4.4 shows an error-bar plot of peak signal-to-noise ratio (PSNR) and bitrate to examine the effect of block size on the rate-distortion performance of DPCM. For this test, 100 trials of an entire BCS-SPL process on Lenna image at a fixed subrate ( $S = M_B/B^2 = 0.2$ ) and quantizer bit depth ( $b = 5$ )<sup>1</sup> are carried out while averages and standard deviations of reconstruction results, such as PSNR and rate, at each block size are obtained. As seen, PSNR results yield a concave graph, with block sizes 4 or 8 giving overall good results. Even though block size 2 yields very high quality sometimes, it is inconsistent and gives a standard deviation as large as about 12 dB. On the other hand, rate results draw a convex graph with block size 4 yielding the lowest bitrate. Also, efficiencies for several test images are calculated in Figure 4.5 by dividing average PSNR by average rate. Higher values means better results in the graphs. As we can observe, block sizes 4 and 8 are good choices, in contrast to 16 or 32 as used in image/video reconstruction without a quantization.

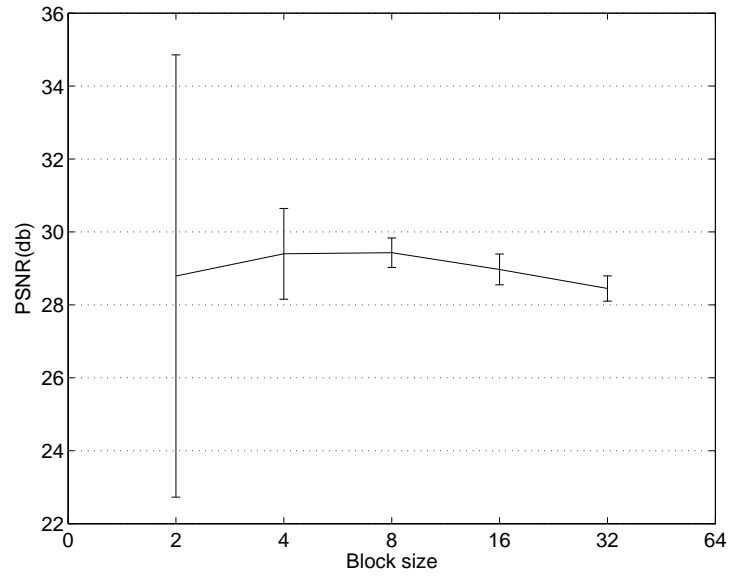
We repeat this block-size-choice process for the other variants, MS-BCS-SPL and MH-BCS-SPL. Going forward, we choose a block size of 8 for BCS-SPL and MH-BCS-SPL, and a block size of 2 for MS-BCS-SPL.

#### 4.2.2 Coding Gain of Quantized BCS-SPL

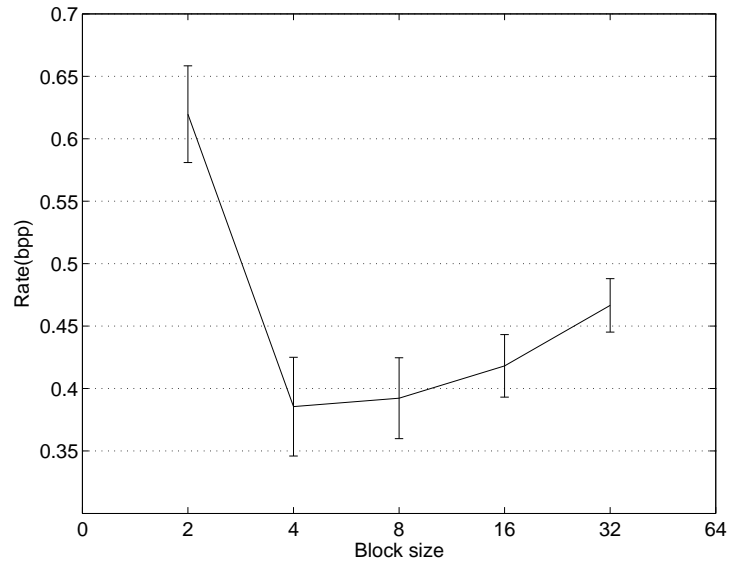
As shown in Sections 4.1.2 and 4.1.3, DPCM can achieve coding gain by reducing the variance of the input signal when the signal is highly correlated, and BCS provides such correlations between blocks. One can measure the rate reduction achieved by DPCM in

---

<sup>1</sup>A quantizer bit depth of  $b$  means that the uniform scalar quantizer has  $2^b$  quantizer bins.



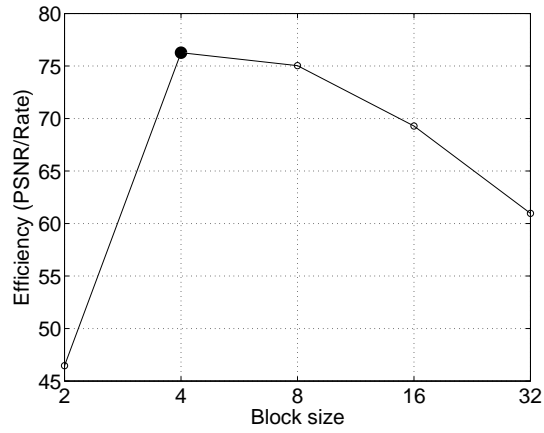
(a)



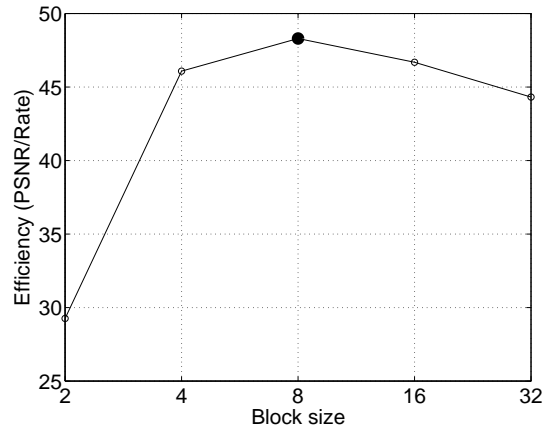
(b)

Figure 4.4

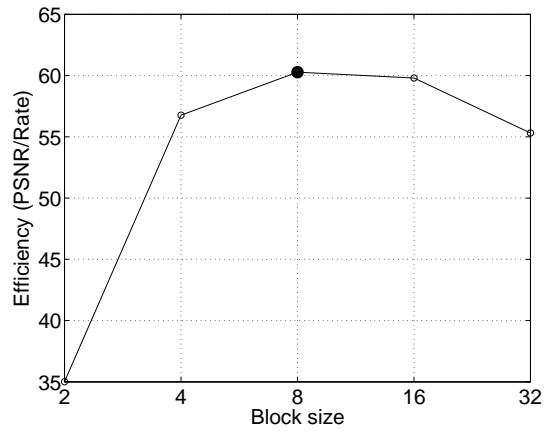
Error-bar plot with respect to block sizes; (a) PSNR (dB), (b) rate (bpp). The averages of 100 trials are observed. Test image is Lenna. Subrate is  $S = 0.2$ , and quantizer bit depth is  $b = 5$ .



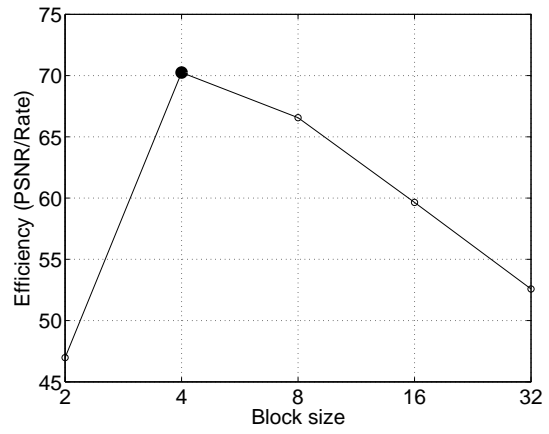
(a) Lenna



(b) Goldhill



(c) Goldhill



(d) Peppers

Figure 4.5

Plot of RD efficiency (PSNR/rate) on various block sizes. Subrate is  $S = 0.2$ , and quantizer bit depth is  $b = 5$ . Black dots indicate maximum values on the graphs.

BCS using (4.14). However, because of the dimensionality reduction in CS, the subrate, which is the ratio between the total number of pixels of the image to be reconstructed and the number of measurements, should also be taken into account. By incorporating the subrate,  $M/N$ , (4.14) can be reformulated as

$$\Delta R = R_{\text{SQ}} - R_{\text{DPCM}} = \frac{1}{2} \frac{M}{N} \log_2 \frac{\sigma_x^2}{\sigma_d^2}. \quad (4.22)$$

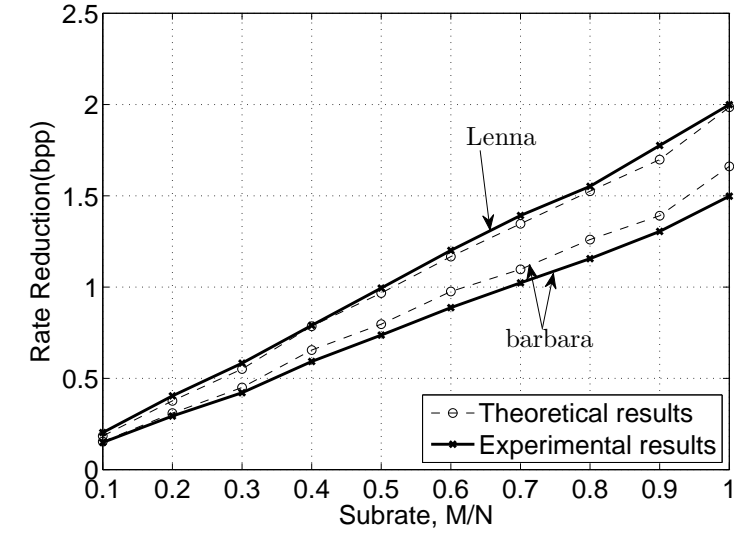
If we use the first-order prediction as used in 1D-DPCM, we have

$$\Delta R = \frac{1}{2} \frac{M}{N} \log_2 \frac{1}{(1 - \rho_1^2)}. \quad (4.23)$$

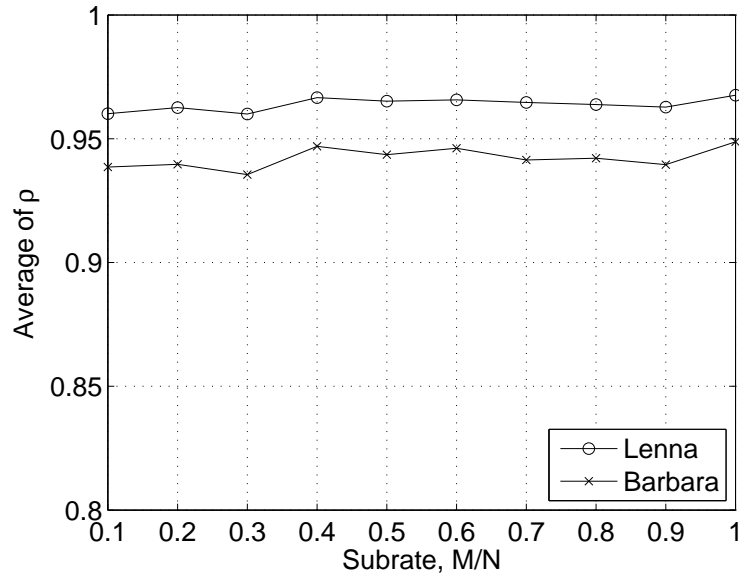
Figure 4.6(a) plots the prediction gain as subrate varies as dictated by (4.23) along with the actual rate reduction as seen experimentally for the  $512 \times 512$  Lenna and Barbara images. For this figure, the theoretical curves use the For the experimental curves, the rate is calculated as the entropy of the measurements. Additionally, Figure 4.6(b) depicts the average correlation as subrate is varied.

As seen in Figure 4.6(a), the reduction in rate is proportional to the subrate. For example, at 0.5 subrate, Lenna ( $\rho_1 = 0.97$ ) has about a 1-bpp prediction gain over uniform SQ, while Barbara ( $\rho_1 = 0.94$ ) has about a 0.75-bpp gain. Barbara has more complex texture and thus lower correlation between blocks than Lenna, resulting a lower DPCM prediction gain.

On the other hand, as seen in Figure 4.6(b), the average correlation does not appear to show any relation to the subrate. This implies that subrate does not have an adverse effect on the correlation of the block measurements, yet the DPCM coding gain increases as the subrate increases. It should be noted that we have observed that the quantizer bit depth



(a)



(b)

Figure 4.6

Prediction gain and average of normalized correlation coefficients.  $512 \times 512$  Lenna and Barbara images are used. Quantization levels are fixed,  $2^b = 2^5 = 32$ . (a) Rate reduction with respect to substrate; (b)  $\bar{\rho}$  with respect to substrate.

does not affect the results much if it is sufficiently large (i.e.,  $b \geq 5$ ); on the other hand, small quantizer bit depths ( $b \leq 3$ ) yield a more significant gap between the theoretical performance anticipated by (4.23) and that obtained experimentally.

### 4.2.3 2D-DPCM for Quantized BCS-SPL

The 1D-DPCM, or line-by-line DPCM, for BCS discussed in the previous section takes advantage of the correlation among blocks in the same horizontal row. However, for images, there exists correlation between current measurement and measurements in other directions as well. Specifically, prediction of a block of measurements is based on the measurements in neighboring blocks, and we propose a third-order prediction such that (4.18) becomes

$$d_m^{(j)} = y_m^{(j)} - \left( a_1 \hat{y}_m^{(j-1)} + a_2 \hat{y}_m^{(j-r)} + a_3 \hat{y}_m^{(j-r-1)} \right), \quad (4.24)$$

where  $r$  is the number of blocks in a row. Consequently, (4.24) forms a prediction of the current block using the block in the same row to the left, the block immediately above in the previous row, and the block diagonally to above left. Even though it is possible to adaptively find the optimal prediction coefficients  $a_1$ ,  $a_2$ , and  $a_3$  to minimize the MSE of the prediction (i.e., using a formulation similar to (4.7) and (4.8)), such an optimization would compromise the simplicity of the encoder with complex prediction-coefficient estimation for each measurement. Instead, we find a set of constants for the coefficients that works well universally (i.e., better than 1D-DPCM) while keeping the encoder as simple as possible.



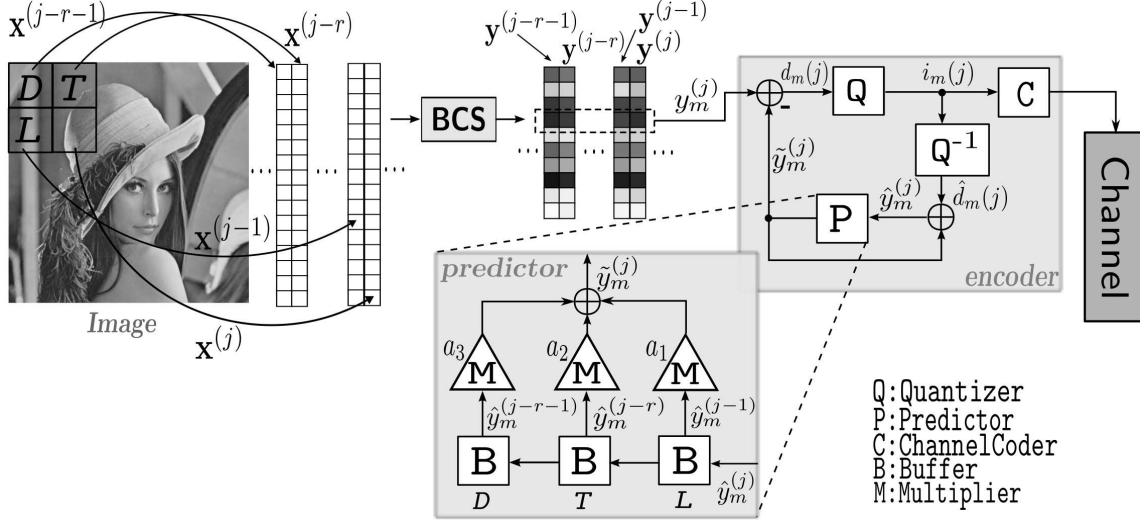


Figure 4.7

Application of 2D-DPCM and SQ to the BCS-SPL architecture. The third-order predictor from (4.24) is used.  $C$  is any entropy coder, such as arithmetic coding.

Figure 4.7 depicts 2D-DPCM applied to the BCS-SPL architecture. Most components are the same as in 1D-DPCM (Figure 4.3), except the predictor that has three buffers (possibly line buffers) to hold the three measurements in the left, top, and diagonal blocks.  $a_1$ ,  $a_2$ , and  $a_3$  are the corresponding constant multipliers of the signals stored in the buffers. Simple summation of the measurements multiplied by constants becomes the prediction of the measurement in the current block.

### 4.3 Experimental Observation

We now present experimental results that demonstrate the performance of the proposed techniques for DPCM-based quantized CS. We first examine the rate-distortion efficiency of 1D- and 2D-DPCM plus SQ by comparing to simple uniform SQ applied alone

to the BCS measurements. We use several BCS-based algorithms—namely, the original BCS-SPL (Section 2.5.1) as well as the MS-BCS-SPL and MH-BCS-SPL extensions from Section 2.6.2 and [24], respectively—and deploy DPCM plus SQ in the framework presented in Figures 4.3 and 4.7 to effectuate quantized CS for all three methods. We note that the implementations of BCS-SPL, MS-BCS-SPL, and MH-BCS-SPL can be found at the BCS-SPL website<sup>2</sup>.

All experiments use  $512 \times 512$  grayscale images, and we measure rate-distortion performance in terms of PSNR in dB and bitrate in bits per pixel (bpp) using the entropy of the quantizer indices as an estimate of the actual bitrate that would be produced by a real entropy coder. The measurement matrix  $\Phi_B$  is an orthonormalized dense Gaussian random matrix, and a 5-level dual-tree discrete wavelet transform (DDWT) [73] is used as the sparsity basis for all three methods. A BCS block size of  $B = 8$  is used for BCS-SPL and MH-BCS-SPL, and MS-BCS-SPL uses  $B = 2$  for each of the levels within the wavelet-based measurement basis. These block sizes are chosen through the process described in Section 4.2.1 and illustrated in Figure 4.5. All SQ is uniform. Finally, we note that, for both SQ as well as DPCM plus SQ, the bitrate obtained depends on both the stepsize of the scalar quantizer as well as the subrate  $S = M_B/B^2$  of the BCS measurement process. The stepsize,  $q$ , is determined by the quantier bit depth,  $b$ , such that

$$q = (\mathbf{y}_{max} - \mathbf{y}_{min})/2^b, \quad (4.25)$$

where  $\mathbf{y}_{max}$  and  $\mathbf{y}_{min}$  are the maximum and minimum value of the measurements, respectively. In all cases, for the experiments here, the optimal combination of quantizer stepsize

---

<sup>2</sup><http://www.ece.msstate.edu/~fowler/BCSSPL/>

Table 4.1

PSNR Performance in dB for a bitrate of 0.5 bpp

<i>Image</i>	BCS-SPL			MS-BCS-SPL			MH-BCS-SPL		
	SQ	DPCM	<i>Gain</i>	SQ	DPCM	<i>Gain</i>	SQ	DPCM	<i>Gain</i>
Lenna	28.1	30.6	+2.5	33.9	34.7	+0.9	29.2	31.4	+2.3
Barbara	23.1	24.2	+1.1	26.6	27.4	+0.8	24.4	27.9	+3.5
Peppers	28.4	30.5	+2.1	33.8	34	+0.2	29.2	31.2	+2.1
Goldhill	26.8	28.1	+1.3	30.6	31	+0.5	26.8	28.8	+2
Man	26.2	27.6	+1.4	30.5	30.7	+0.2	26.5	27.9	+1.4
Clown	27.2	28.8	+1.6	32.7	33.2	+0.5	28.4	30.8	+2.4
<i>Average</i>	26.6	28.3	+1.7	31.3	31.8	+0.5	27.4	29.7	+2.3

and subrate is chosen via an exhaustive search over all possible (stepsize, subrate) pairs drawn from a finite set of stepsizes and a finite set of subrates. Figure 4.8 illustrates this exhaustive-search process via a convex rate-distortion hull.

#### 4.3.1 1D-BCS-SPL-DPCM with SQ

In this section, we examine 1D-DPCM plus SQ applied to the BCS-SPL architecture as depicted in Figure 4.3.

We compare to use of SQ alone. Table 4.1 compares the PSNR performance at a fixed bitrate of 0.5 bpp for the three BCS-based techniques, BCS-SPL, MS-BCS-SPL, and MH-BCS-SPL. We see that, for all three algorithms, the addition of DPCM to the quantization process increases the PSNR by 0.5 to 2.3 dB on average as compared to simply using SQ alone. MS-BCS-SPL has lower PSNR gain than BCS-SPL or MH-BCS-SPL because only the baseband of the MS-BCS-SPL possess the correlation between block measurements. It is observed that the discrete wavelet transform (DWT) within the measurement process

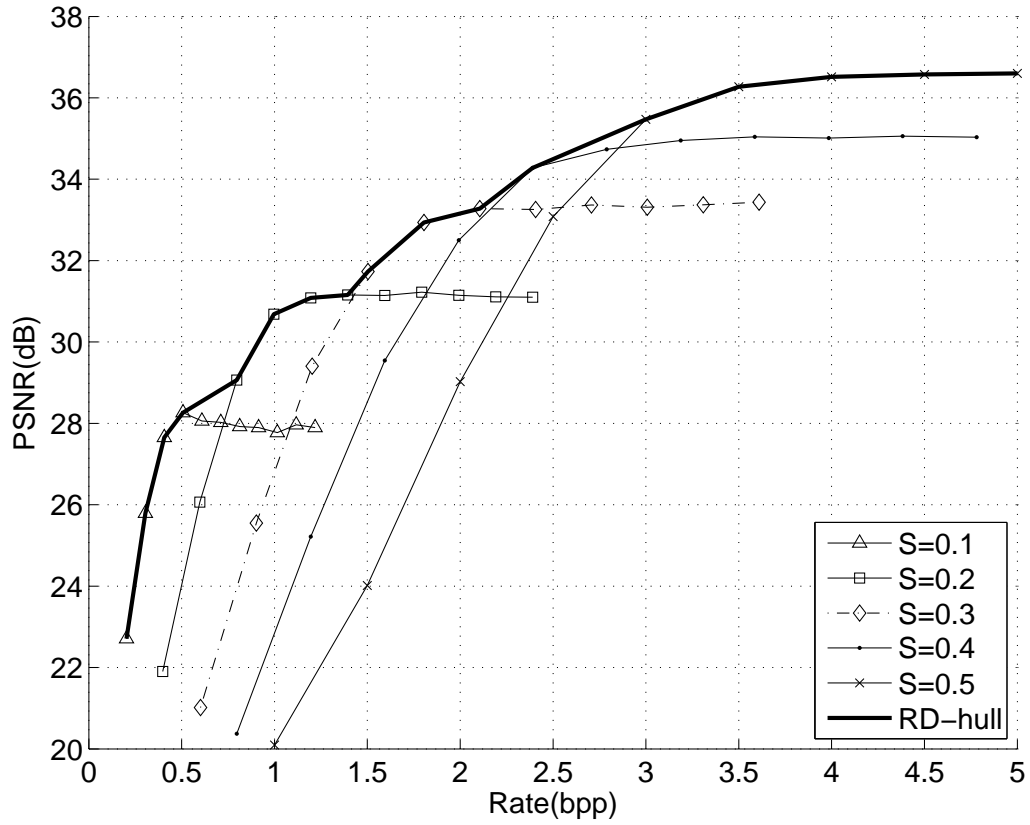


Figure 4.8

The convex rate-distortion hull. Rate-distortion curves are drawn for various subrates  $S$  by varying the quantizer bit depth  $b$ . The convex rate-distortion hull is then the convex hull of all curves. The convex rate-distortion hull then yields the specific (stepsize, subrate) pair to use to obtain maximal PSNR at a specific bitrate.

of MS-BCS-SPL decorrelates the coefficients in subbands other than the baseband; as a consequence, the random projected subbands do not show high correlation. In the case of MH-BCS-SPL, the gain over the SQ-alone result is significant since MH-BCS-SPL uses an initial BCS-SPL reconstruction to find the multihypothesis weights (see [24]; DPCM provides better quality for this initial reconstruction for same bitrate.

Figures 4.9–4.11 illustrate the corresponding rate-distortion performance. In general, all three algorithms using DPCM plus SQ outperform their counterparts using SQ alone. As explained above, the gain due to DPCM for MS-BCS-SPL is not as significant as it is for the other two algorithms since DPCM is applied to only the baseband in MS-BCS-SPL.

### 4.3.2 2D-BCS-SPL-DPCM with SQ

In this section, we examine 2D-DPCM plus SQ applied to the BCS-SPL architecture in Figure 4.3. The results are compared to the corresponding 1D-DPCM plus SQ. Rate-distortion performance is observed with several sets of prediction coefficients— $a_1$ ,  $a_2$ , and  $a_3$ —that sum 1.

In order to decide the optimal prediction coefficients non-adaptively, we perform BCS-SPL with 2D-DPCM using several sets of prediction coefficients. The results are shown in Figures 4.12–4.15. As can be observed, the prediction coefficients for the 2D-DPCM do not significant affect the performance. Overall, the best performance using 2D-DPCM is obtained for  $a_1 = a_2 = 0.5$  and  $a_3 = 0$ . Therefore, we will use this set of prediction coefficients for 2D-DPCM hereafter.

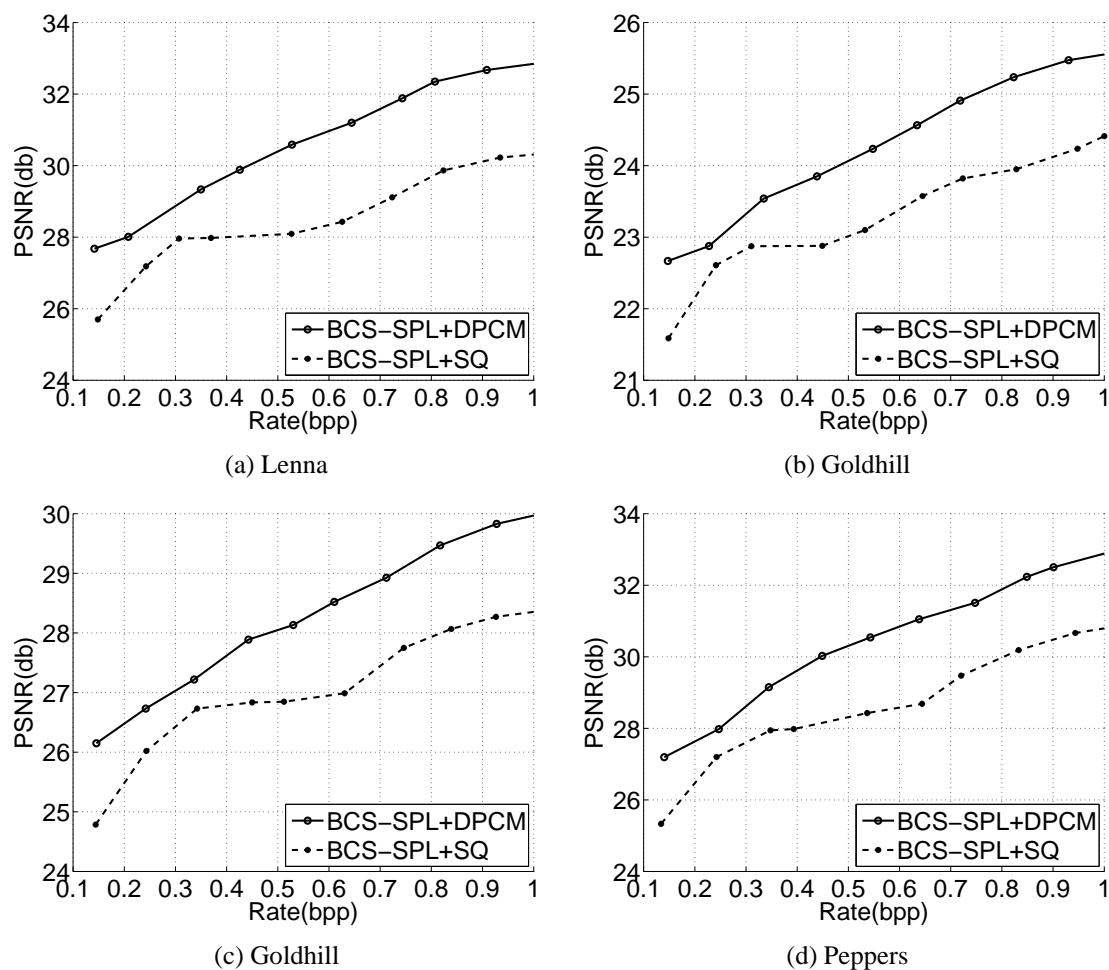


Figure 4.9

Rate-distortion performance of 1D-DPCM plus SQ applied to BCS-SPL.

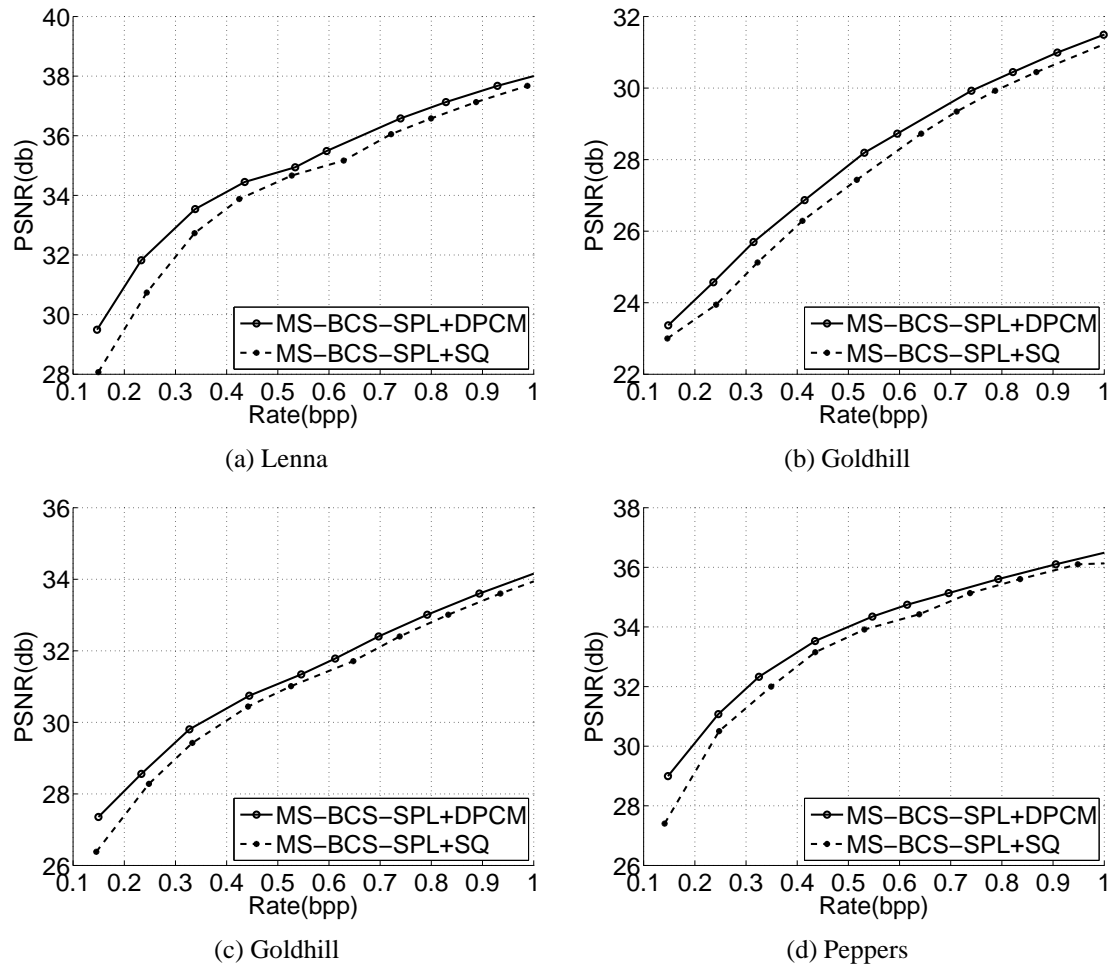
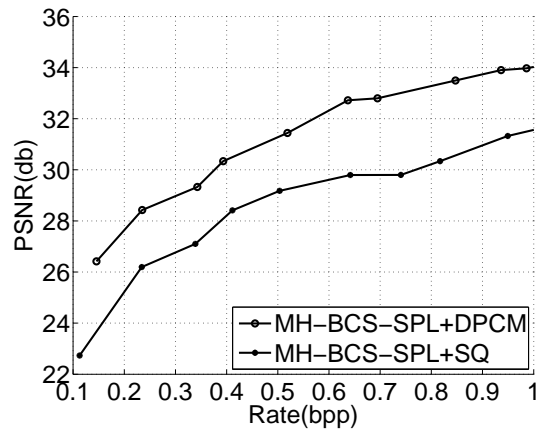
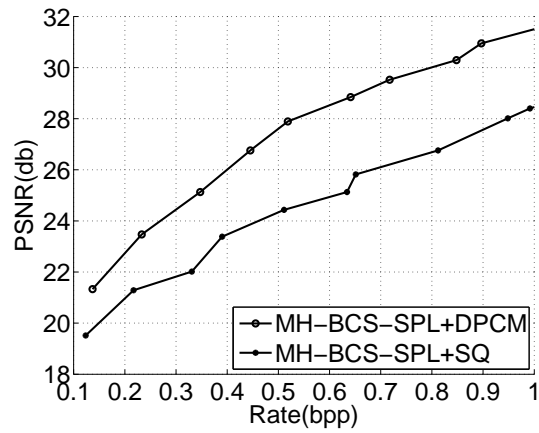


Figure 4.10

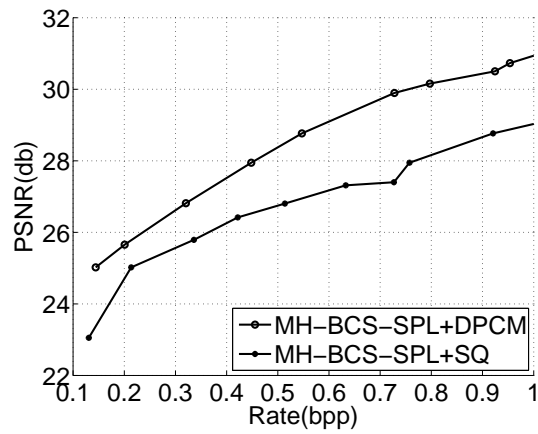
Rate-distortion performance of 1D-DPCM plus SQ applied to MS-BCS-SPL.



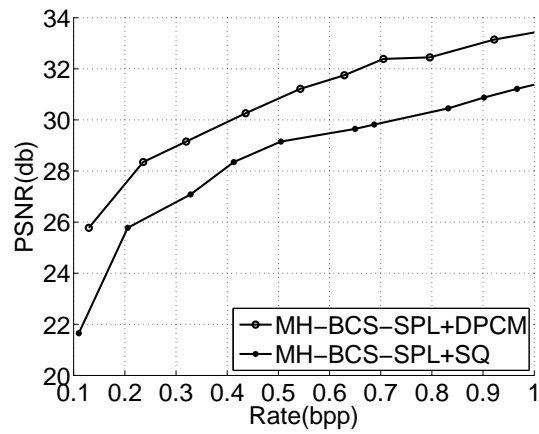
(a) Lenna



(b) Goldhill



(c) Goldhill



(d) Peppers

Figure 4.11

Rate-distortion performance of 1D-DPCM plus SQ applied to MH-BCS-SPL.



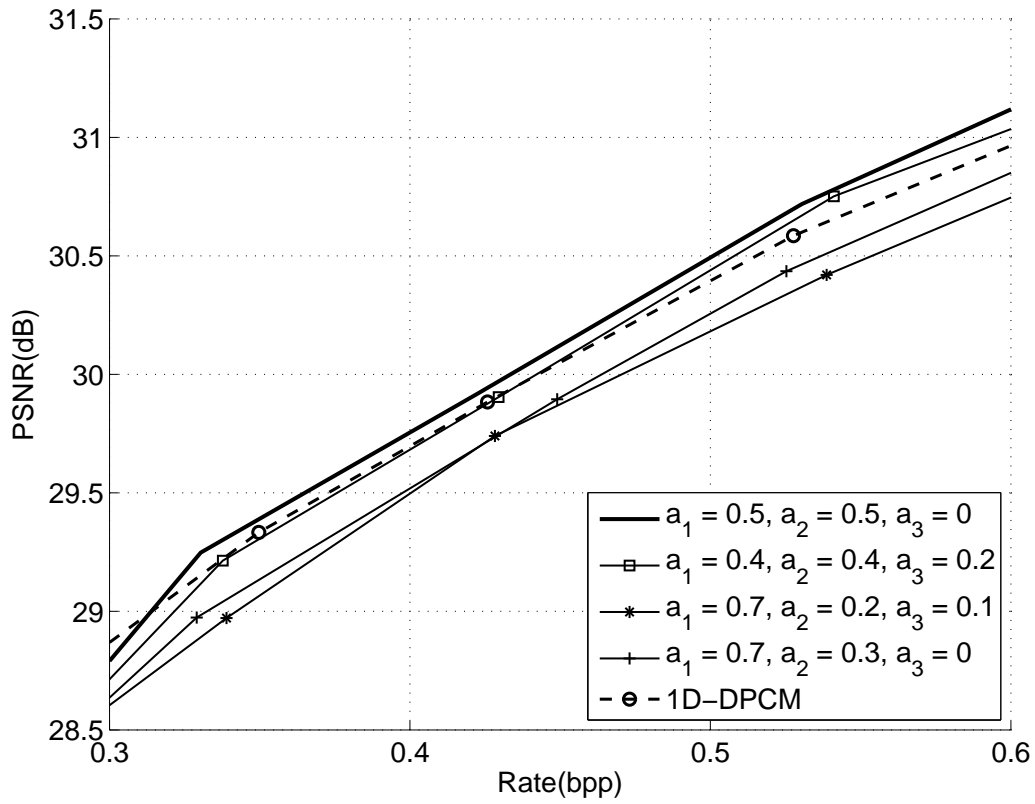


Figure 4.12

RD performance of various predictions in 2D-DPCM for Lenna.

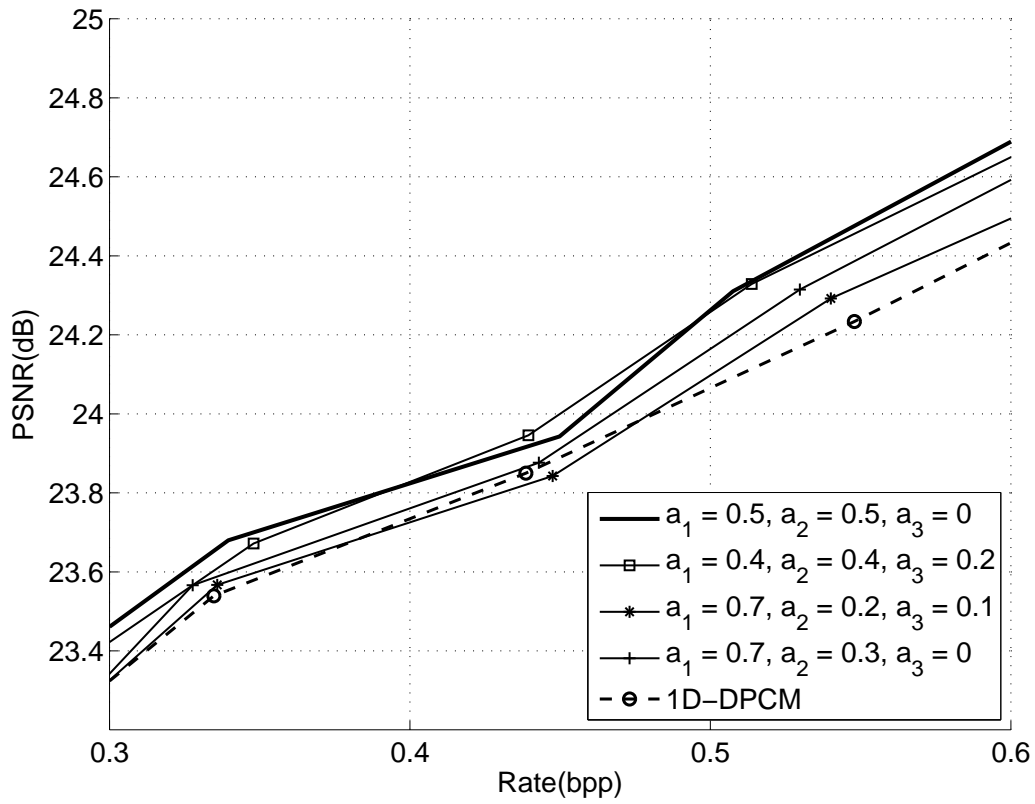


Figure 4.13

RD performance of various prediction coefficients in 2D-DPCM for Barbara.

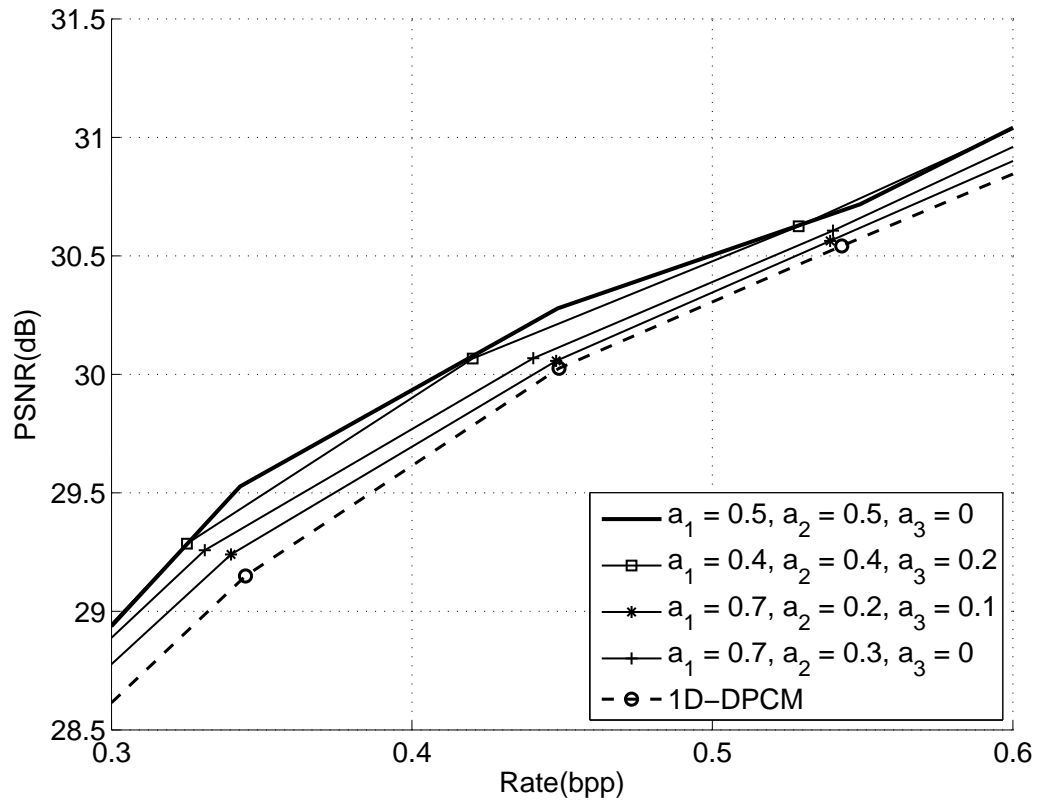


Figure 4.14

RD performance of various prediction coefficients in 2D-DPCM for Peppers.

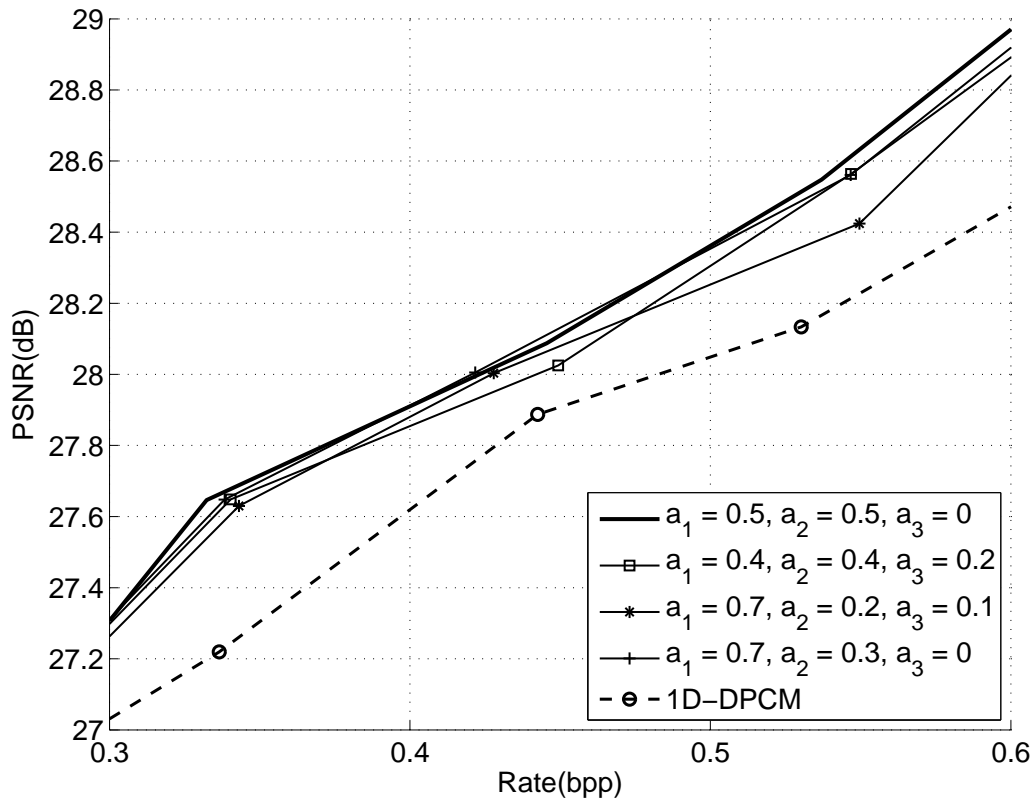


Figure 4.15

RD performance of various prediction coefficients in 2D-DPCM for Goldhill.

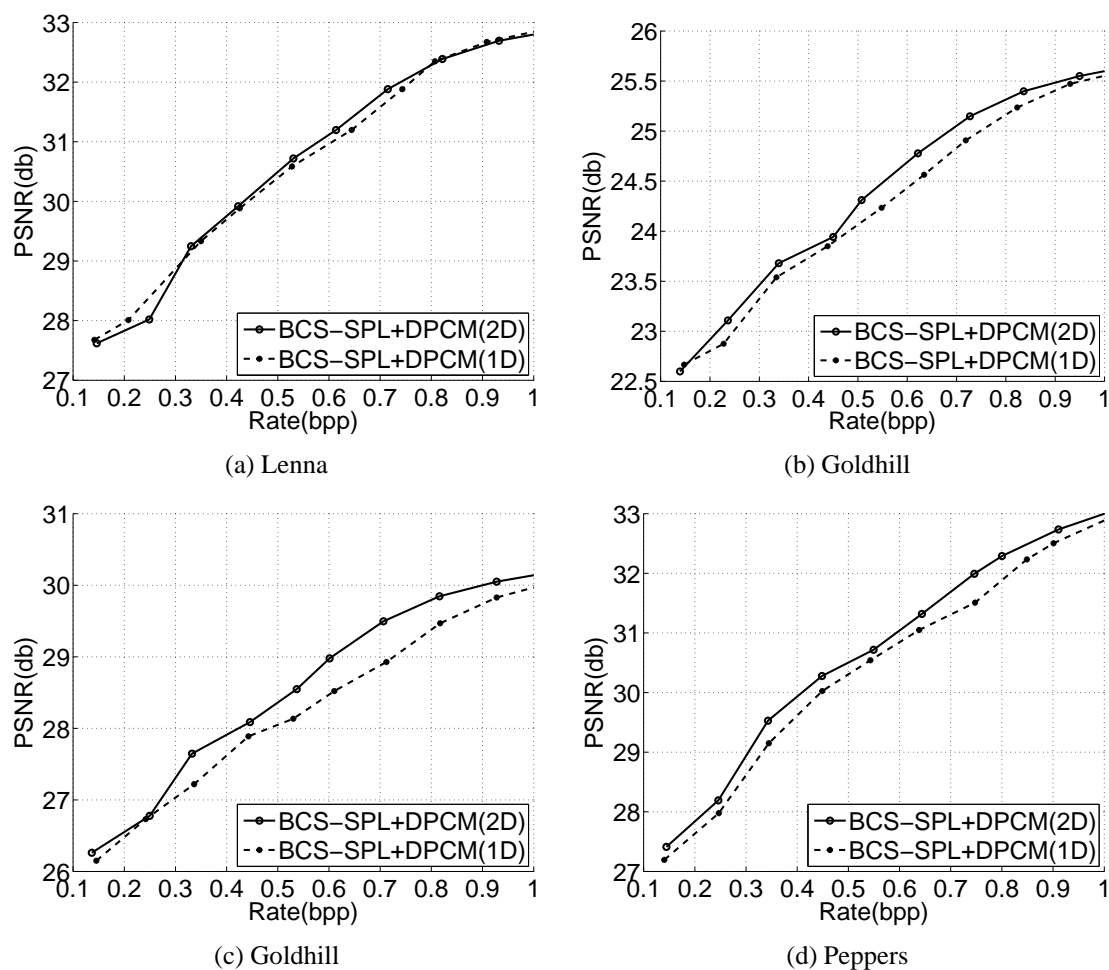
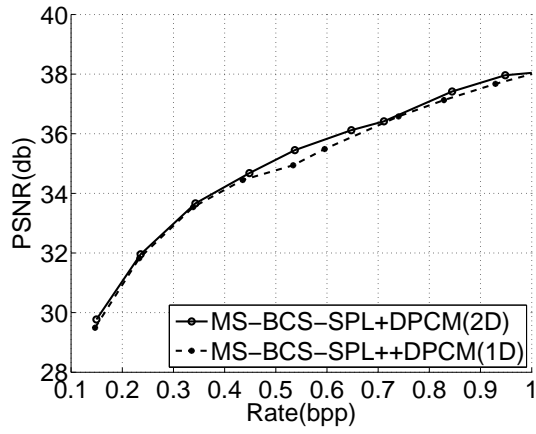
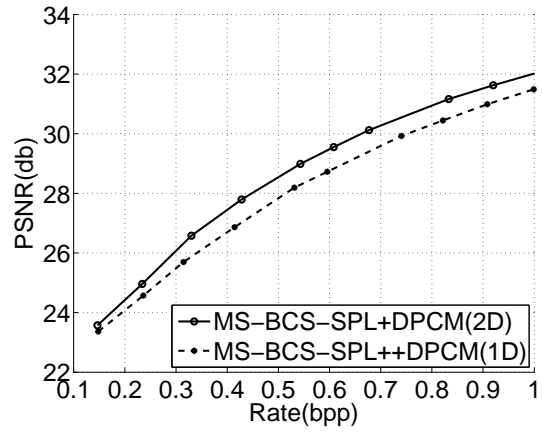


Figure 4.16

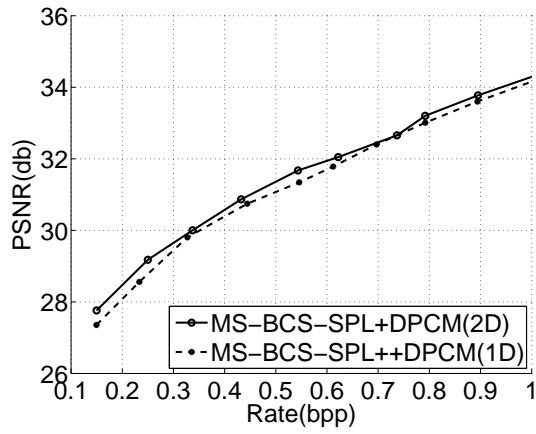
RD performance of 2D-DPCM plus SQ applied to BCS-SPL.



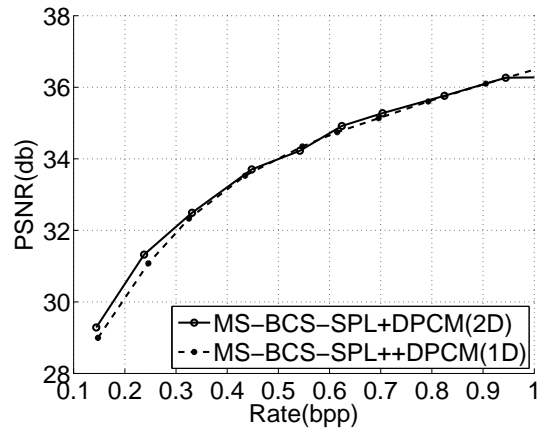
(a) Lenna



(b) Goldhill



(c) Goldhill



(d) Peppers

Figure 4.17

RD performance of 2D-DPCM plus SQ applied to MS-BCS-SPL.

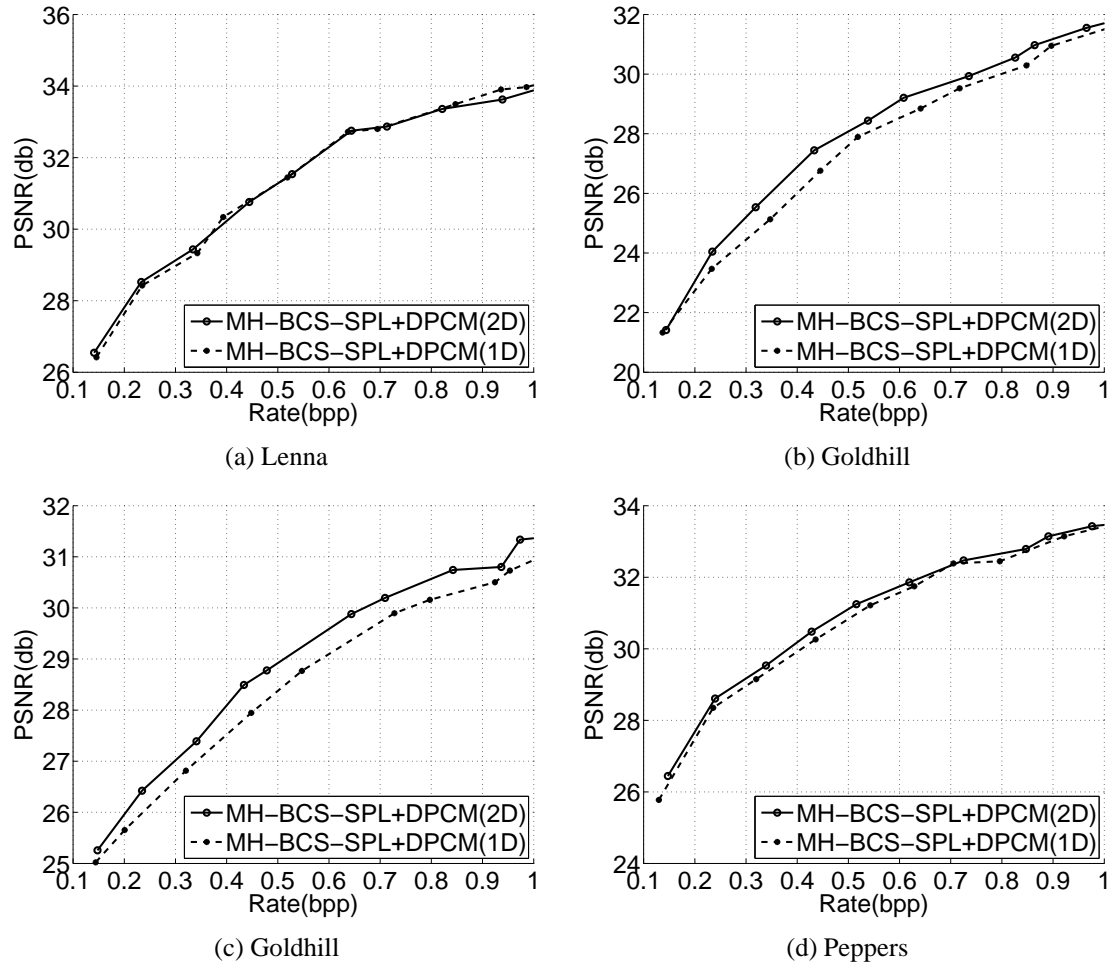


Figure 4.18

RD performance of 2D-DPCM plus SQ applied to MH-BCS-SPL.

Figures 4.16–4.18 depict the rate-distortion performance using 2D-DPCM plus SQ applied to BCS-SPL, MS-BCS-SPL, and MH-BCS-SPL, respectively. The results are also compared to the 1D-DPCM variant. The PSNR gain over 1D-DPCM is not as great as was achieved for 1D-DPCM with respect to the use of SQ alone, but some improvement for some images can be found in the results. Ultimately, while the 2D prediction does not degrade the performance, it is not clear that the modest coding gain merits the corresponding increase in complexity arising from the 2D prediction.

### 4.3.3 Comparison of Various CS Techniques for Quantized Measurements

We now compare all three BCS-based techniques with other quantized CS reconstruction approaches. Because 2D-DPCM performs better than 1D, only 2D-DPCM results are compared in this section. 1D-DPCM results can be found in [84].

Figures 4.19–4.22 present the rate-distortion performance for all three BCS-based techniques using the 2D-DPCM plus SQ framework for a bitrate ranging from 0.1 to 1.5 bpp. Additionally, Figures 4.19–4.22 include as benchmarks two other quantized-CS approaches—model-based adaptive recovery of compressive sensing (MARX) [111] using the PQ proposed in [108] and BPDQ<sup>3</sup> [65]. Finally, we also include the rate-distortion performance of JPEG as indicative of the performance of a relatively simple image coder built with traditional source-coding techniques. Generally, we see that the DPCM-based BCS-SPL and MH-BCS-SPL match the performance of other quantized-CS techniques, while MS-BCS-SPL reconstruction outperforms the others. However, traditional source-coding in the form

---

<sup>3</sup><http://wiki.epfl.ch/bpdq>



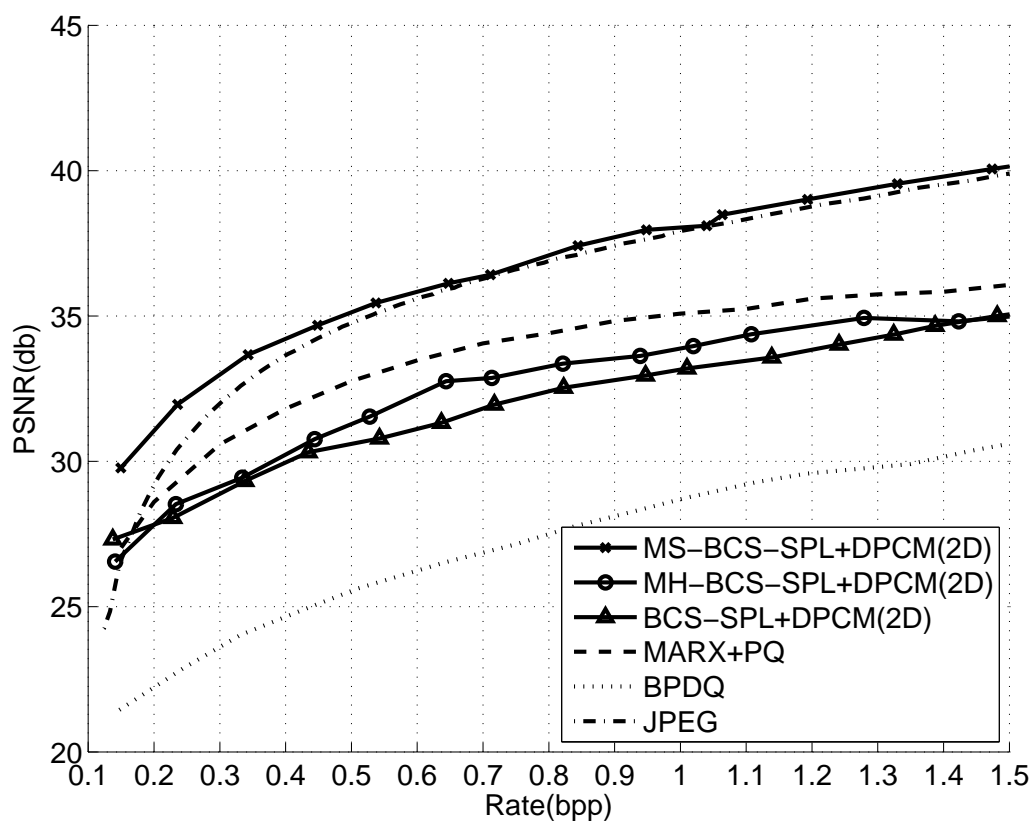


Figure 4.19

Rate-distortion performance for Lenna

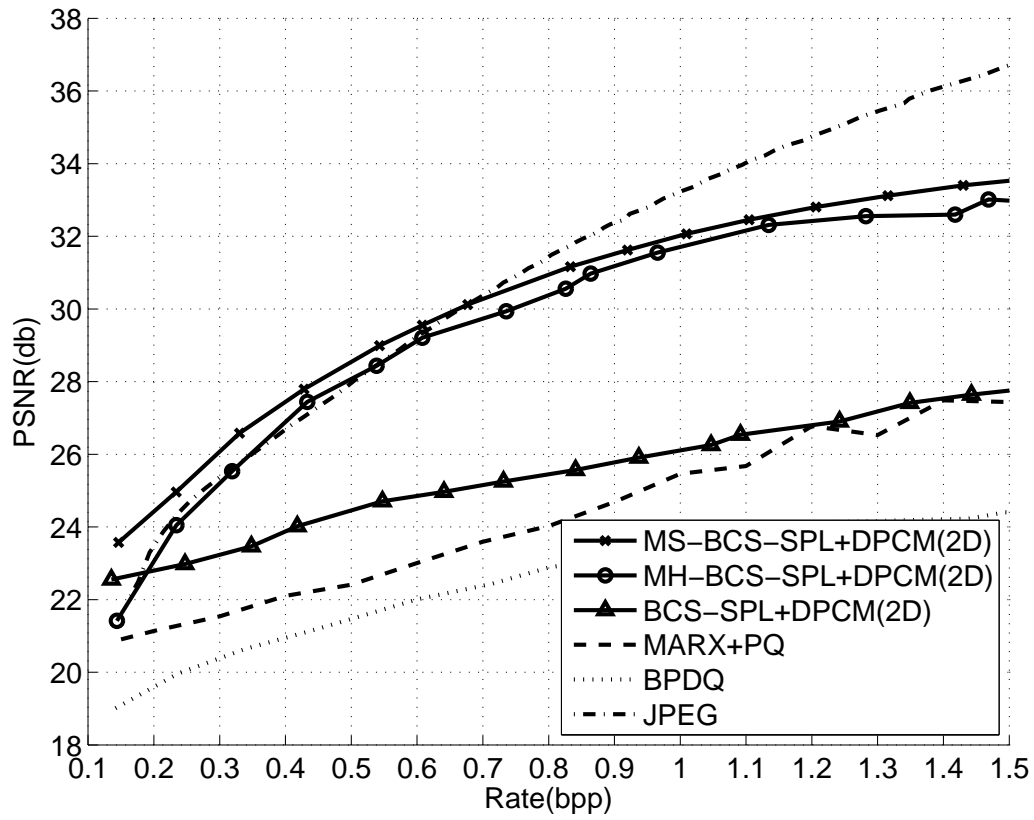


Figure 4.20

Rate-distortion performance for Barbara

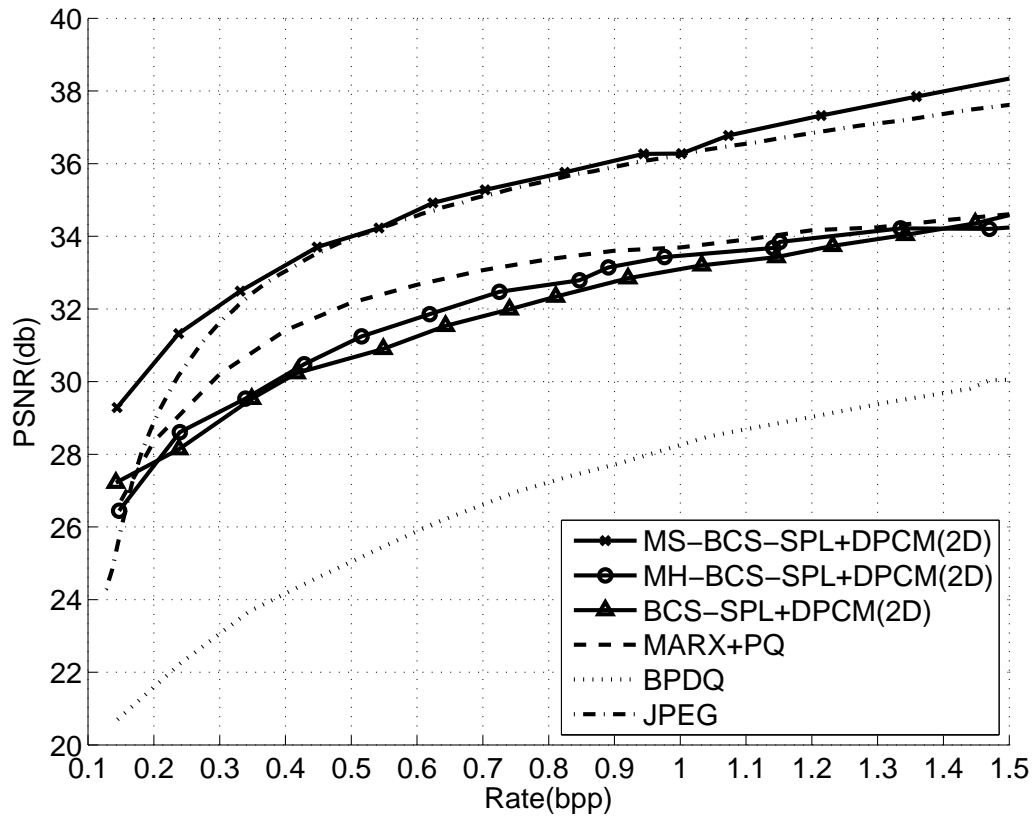


Figure 4.21

Rate-distortion performance for Peppers

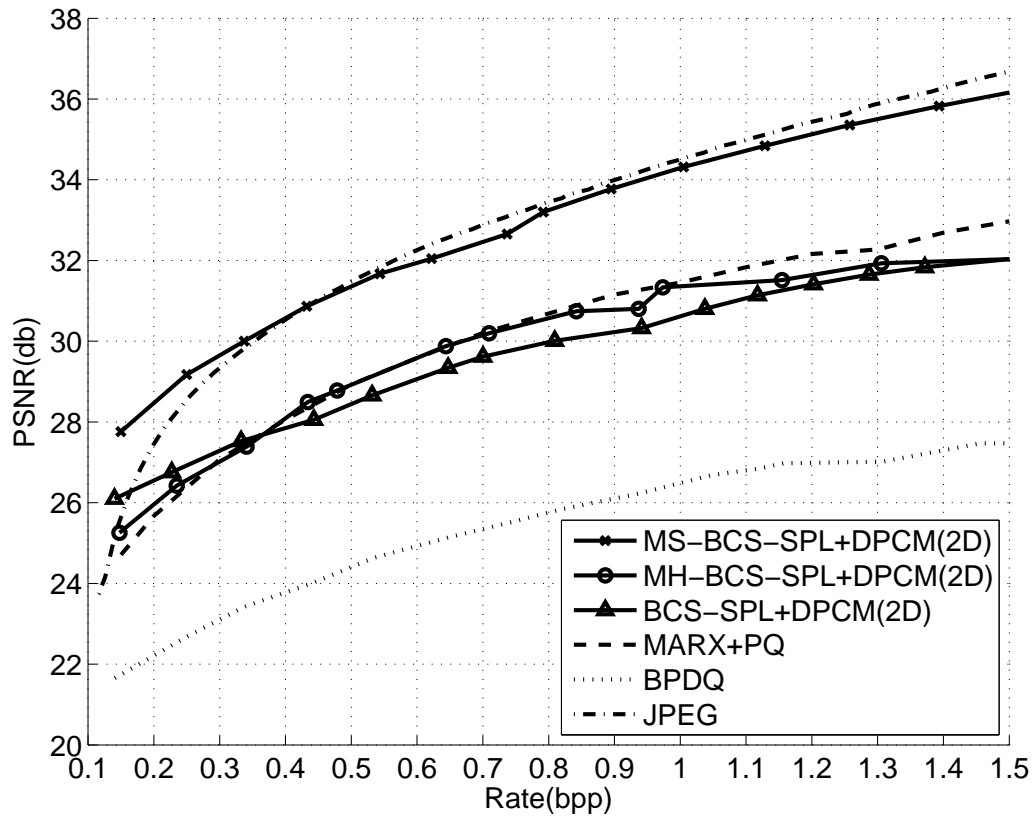


Figure 4.22

Rate-distortion performance for Goldhill

of JPEG achieves the best rate-distortion performance, except at low bitrates (0.5 bpp and below) where MS-BCS-SPL with DPCM yields higher PSNR. Execution time is not much different from that already tabulated in Table 2.4.

Figure 4.23 provides the Lenna image reconstructed at a bitrate of 0.25 bpp; these figures show only a detailed portion from the center of the image. We can see that MS-BCS-SPL with 2D-DPCM provides better overall visual quality with sharper contours than other techniques.



(a) MS-BCS-SPL+2D-DPCM, 31.8 dB



(b) JPEG, 30.9 dB



(c) MARX+PQ, 29.8 dB



(d) MH-BCS-SPL+2D-DPCM, 28.5 dB



(e) BCS-SPL+2D-DPCM, 28.5 dB



(f) BPDQ, 23.1 dB

Figure 4.23

Reconstructions of the  $512 \times 512$  “Lenna” image (shown in detail) for a bitrate of 0.25 bpp.

#### 4.4 Remarks

In this chapter, we proposed the incorporation of DPCM to achieve quantized CS of images based on blocks. In essence, we used one measurement-domain block to predict the next, applying uniform SQ to the measurement-domain residual of the prediction. Experimental results demonstrated an improvement of 0.5 to 2 dB in rate-distortion performance as compared to BCS-based image reconstruction using uniform SQ alone. Additionally, rate-distortion performance superior to alternative quantized-CS schemes relying on optimized quantization or reconstruction was observed.

A key benefit of our proposed DPCM-based methodology is that both the BCS-based sensor as well as the BCS-based reconstruction are unmodified; in fact, the latter can be any BCS-based reconstruction. While the sensor device does incur some additional complexity, the addition of the 1D-DPCM processing (a subtraction) is not substantially more burdensome than the already-necessary SQ. 2D-DPCM, with additional cost of buffers and multipliers, yields a modest coding gain with respect to 1D-DPCM for some images. Experimental results using state-of-the-art BCS-based reconstruction algorithms on still images demonstrate that, not only does this simple DPCM-plus-SQ approach to quantized CS provide rate-distortion performance surprisingly competitive with that of alternative approaches such as [65, 109], it can occasionally rival traditional image coding in the form of JPEG, particularly at low bitrates.

## CHAPTER 5

### CONCLUSIONS

In this dissertation, the emerging concept of compressed sensing (CS) has been studied with a particular focus on recent proposals for its use with a variety of imaging media, including still images and motion video. Firstly, we have addressed the CS image reconstruction problem which is a challenging task because the computational complexity grows quickly as the size of problem increases due to the non-linear reconstruction process and a huge measurement matrix. By splitting the problem into small blocks, burdens of storing and handling the huge measurement matrix are effectively eased. The projected Landweber approach to CS reconstruction coupled with directional transforms and statistical thresholding provides high-quality image reconstruction without huge computational load in the form of the block-based compressed sensing with smooth projected Landweber reconstruction (BCS-SPL) algorithm. As an extension, multiscale BCS-SPL (MS-BCS-SPL) deploys BCS-SPL in the wavelet domain to provide multiscale measurement and reconstruction. The performance of our proposed block-based compressed sensing (BCS)-based techniques with several other prominent CS reconstruction proposed in recent literature were observed. We found that BCS-based techniques are competitive with other approaches relying on a full-frame CS measurement operator yet take advantage of significantly reduced memory and computation.



Secondly, BCS-SPL can be adapted to the CS video reconstruction problem which is even more complex because of the data size. We have capitalized on established strategies in traditional source coding for motion estimation and compensation. This predictive technique was used to exploit frame-to-frame redundancies due to object motion to provide residual frames that are more compressible—in a sense of a rapid decay of coefficient magnitude within a sparsity transform—than their corresponding original frames. Reconstruction from such prediction residuals was seen to significantly improve performance as compared to straightforward reconstruction of each frame independently as well as a 3D joint reconstruction that features temporal decorrelation in the form of a 3D transform, but does not exploit motion. Experimental results show that the motion-compensated version of BCS-SPL achieves performance on par with, and sometimes superior to, other CS-video approaches.

Finally, we have addressed the issue of quantization and CS for images. In a situation wherein analog hardware acquires images, the CS measurements are real-valued. Yet, the reconstruction procedures that follow acquisition are necessarily implemented within digital computers. As a consequence, some form of quantization is unavoidable in the conversion of the signal from analog to digital. However, the issue of quantization in conjunction with CS has not been widely considered in the literature until recently. Differential pulse-code modulation (DPCM), commonly used in signal processing, exploits sample-to-sample correlation so that coding efficiency increases (or bits to code the signal decreases) when correlation is high. Such correlation resides in the measurements across the blocks generated by BCS. Therefore, application of DPCM plus scalar quanti-

zation (SQ) in the BCS-SPL architecture achieved significant bit savings as compared to SQ of BCS-SPL. Its extension, DPCM using 2D prediction, also provided gain over 1D prediction. Three BCS-based algorithms were employed to investigate the rate-distortion performance DPCM, comparing to other quantized CS techniques. We found the BCS-SPL algorithms with DPCM to be on par with, or superior to, alternative quantized CS techniques as well as a traditional source coding on the form of JPEG.

## REFERENCES

- [1] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Fast Image Recovery Using Variable Splitting and Constrained Optimization,” *IEEE Transactions on Information Theory*, vol. 19, no. 9, September 2010, pp. 2345–2356.
- [2] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems,” *IEEE Transactions on Information Theory*, vol. 20, no. 3, March 2011, pp. 681–695.
- [3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image Coding Using Wavelet Transform,” *IEEE Transactions on Image Processing*, vol. 1, no. 2, April 1992, pp. 205–220.
- [4] A. Averbuch, S. Dekel, and S. Deutsch, “Adaptive Compressed Image Sensing Using Dictionaries,” *SIAM J. Img. Sci.*, vol. 5, no. 1, January 2012, pp. 57–89.
- [5] S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Bayesian Compressive Sensing Using Laplace Priors,” *IEEE Transactions on Image Processing*, vol. 19, no. 1, January 2010, pp. 53–63.
- [6] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, “A Simple Proof of the Restricted Isometry Property for Random Matrices,” *Constructive Approximation*, vol. 28, no. 3, December 2008, pp. 253–263.
- [7] R. G. Baraniuk, E. Candès, R. Nowak, and M. Vetterli, “Compressive Sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, March 2008, pp. 12–13.
- [8] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-Based Compressive Sensing,” *IEEE Transactions on Information Theory*, vol. 56, no. 4, April 2010, pp. 1982–2001.
- [9] R. G. Baraniuk, V. Cevher, and M. B. Wakin, “Low-Dimensional Models for Dimensionality Reduction and Signal Recovery: A Geometric Perspective,” *Proceedings of the IEEE*, vol. 98, no. 6, June 2010, pp. 959–971.
- [10] A. Beck and M. Teboulle, “A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, 2009, pp. 183–202.

- [11] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, Bristol, UK, 1998.
- [12] J. M. Bioucas-Dias and M. A. T. Figueiredo, “A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration,” *IEEE Transactions on Image Processing*, vol. 16, no. 12, December 2007, pp. 2992–3004.
- [13] T. Blumensath and M. E. Davies, “Iterative Thresholding for Sparse Approximations,” *The Journal of Fourier Analysis and Applications*, vol. 14, no. 5, December 2008, pp. 629–654.
- [14] T. Blumensath and M. E. Davies, “Iterative Hard Thresholding for Compressed Sensing,” *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, November 2009, pp. 265–274.
- [15] J. B. Boettcher and J. E. Fowler, “Video Coding Using a Complex Wavelet Transform and Set Partitioning,” *IEEE Signal Processing Letters*, vol. 14, no. 9, September 2007, pp. 633–636.
- [16] P. T. Boufounos and R. G. Baraniuk, “1-Bit Compressive Sensing,” *Proceedings of the 42<sup>nd</sup> Annual Conference on Information Sciences and Systems*, Princeton, NJ, March 2008, pp. 16–21.
- [17] E. Candès, J. Romberg, and T. Tao, “Stable Signal Recovery from Incomplete and Inaccurate Measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, August 2006, pp. 1207–1223.
- [18] E. Candès and T. Tao, “Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies?,” *IEEE Transactions on Information Theory*, vol. 52, no. 12, December 2006, pp. 5406–5425.
- [19] E. Candès, M. B. Wakin, and S. P. Boyd, “Enhancing Sparsity by Reweighted  $\ell_1$  Minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, December 2008, pp. 877–905.
- [20] E. J. Candès, “Compressive Sampling,” *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, August 2006, vol. 3, pp. 1433–1452.
- [21] E. J. Candès and T. Tao, “Decoding by Linear Programming,” *IEEE Transactions on Information Theory*, vol. 51, no. 12, December 2005, pp. 4203–4215.
- [22] E. J. Candès and M. B. Wakin, “An Introduction To Compressive Sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, March 2008, pp. 21–30.
- [23] V. Chappelier, C. Guillemot, and S. Marinković, “Image Coding with Iterated Contourlet and Wavelet Transforms,” *Proceedings of the International Conference on Image Processing*, Singapore, October 2004, vol. 5, pp. 3157–3160.

- [24] C. Chen, E. W. Tramel, and J. E. Fowler, “Compressed-Sensing Recovery of Images and Video Using Multihypothesis Predictions,” *Proceedings of the 45<sup>th</sup> Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2011, pp. 1193–1198.
- [25] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic Decomposition by Basis Pursuit,” *SIAM Journal on Scientific Computing*, vol. 20, no. 1, August 1998, pp. 33–61.
- [26] L. Şendur and I. W. Selesnick, “Bivariate Shrinkage Functions for Wavelet-Based Denoising Exploiting Interscale Dependency,” *IEEE Transactions on Signal Processing*, vol. 50, no. 11, November 2002, pp. 2744–2756.
- [27] W. Dai, H. V. Pham, and O. Milenkovic, “Distortion-Rate Functions for Quantized Compressive Sensing,” *IEEE Information Theory Workshop on Networking and Information Theory*, Volos, Greece, June 2009, pp. 171–175.
- [28] I. Daubechies, M. Defrise, and C. De Mol, “An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constraint,” *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, November 2004, pp. 1413–1457.
- [29] I. Daubechies, R. DeVore, M. Fournasier, and C. S. Güntürk, “Iteratively Reweighted Least Squares Minimization for Sparse Recovery,” *Communications on Pure and Applied Mathematics*, vol. 63, no. 1, January 2010, pp. 1–38.
- [30] I. Daubechies and W. Sweldens, “Factoring Wavelet Transforms into Lifting Steps,” *The Journal of Fourier Analysis and Applications*, vol. 4, no. 3, 1998, pp. 245–267.
- [31] M. A. Davenport, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, *A Simple Proof that Random Matrices are Democratic*, Tech. Rep. TREE-0906, Rice University ECE Department, November 2009.
- [32] M. N. Do and M. Vetterli, “The Contourlet Transform: An Efficient Directional Multiresolution Image Representation,” *IEEE Transactions on Image Processing*, vol. 14, no. 12, December 2005, pp. 2091–2106.
- [33] T. T. Do, Y. Chen, D. T. Nguyen, N. Nguyen, L. Gan, and T. D. Tran, “Distributed Compressed Video Sensing,” *Proceedings of the International Conference on Image Processing*, Cairo, Egypt, November 2009, pp. 1393–1396.
- [34] T. T. Do, L. Gan, N. Nguyen, and T. D. Tran, “Sparsity Adaptive Matching Pursuit Algorithm for Practical Compressed Sensing,” *Proceedings of the 42<sup>th</sup> Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, October 2008, pp. 581–587.

- [35] T. T. Do, T. D. Tran, and L. Gan, "Fast Compressive Sampling with Structurally Random Matrices," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, NV, March 2008, pp. 3369–3372.
- [36] D. L. Donoho, "De-Noising by Soft-Thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, May 1995, pp. 613–627.
- [37] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, April 2006, pp. 1289–1306.
- [38] D. L. Donoho, A. Maleki, and A. Montanari, "Message-Passing Algorithms for Compressed Sensing," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 45, November 2009, pp. 18914–18919.
- [39] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, *Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit*, Tech. Rep., Stanford University, 2006.
- [40] M. F. Duarte and R. G. Baraniuk, "Kronecker Compressive Sensing," *IEEE Transactions on Image Processing*, vol. 21, no. 2, February 2012, pp. 494–504.
- [41] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-Pixel Imaging via Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, March 2008, pp. 83–91.
- [42] M. F. Duarte and Y. C. Eldar, "Structured Compressed Sensing: From Theory to Applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, September 2011, pp. 4053–4085.
- [43] R. Eslami and H. Radha, "On Low Bit-Rate Coding Using the Contourlet Transform," *Proceedings of the 37<sup>th</sup> Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 2003, vol. 2, pp. 1524–1528.
- [44] R. Eslami and H. Radha, "Wavelet-Based Contourlet Transform and its Application to Image Coding," *Proceedings of the International Conference on Image Processing*, Singapore, October 2004, vol. 5, pp. 3189–3192.
- [45] E. Feig and S. Winograd, "Fast Algorithms for the Discrete Cosine Transform," *IEEE Transactions on Signal Processing*, vol. 40, no. 9, September 1992, pp. 2174–2193.
- [46] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems," *IEEE Journal on Selected Areas in Communications*, vol. 1, no. 4, December 2007, pp. 586–597.

- [47] J. E. Fowler, J. B. Boettcher, and B. Pesquet-Popescu, "Image Coding Using a Complex Dual-Tree Wavelet Transform," *Proceedings of the European Signal Processing Conference*, Poznań, Poland, September 2007.
- [48] J. E. Fowler, S. Mun, and E. W. Tramel, "Multiscale Block Compressed Sensing with Smoother Projected Landweber Reconstruction," *Proceedings of the European Signal Processing Conference*, Barcelona, Spain, August 2011, pp. 564–568.
- [49] J. E. Fowler, S. Mun, and E. W. Tramel, "Block-Based Compressed Sensing of Images and Video," *Foundations and Trends in Signal Processing*, vol. 4, no. 4, March 2012, pp. 297–416.
- [50] L. Gan, "Block Compressed Sensing of Natural Images," *Proceedings of the International Conference on Digital Signal Processing*, Cardiff, UK, July 2007, pp. 403–406.
- [51] L. Gan, T. T. Do, and T. D. Tran, "Fast Compressive Imaging Using Scrambled Block Hadamard Ensemble," *Proceedings of the European Signal Processing Conference*, Lausanne, Switzerland, August 2008.
- [52] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Norwell, MA, 1992.
- [53] B. Girod, "Efficiency Analysis of Multihypothesis Motion-Compensated Prediction for Video Coding," *IEEE Transactions on Image Processing*, vol. 9, no. 2, February 2000, pp. 173–183.
- [54] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive Sampling and Lossy Compression," *IEEE Signal Processing Magazine*, vol. 25, no. 2, March 2008, pp. 48–56.
- [55] C. S. Güntürk, M. Lammers, A. Powell, R. Saab, and Ö. Yılmaz, "Sobolev Duals for Random Frames and Sigma-Delta Quantization of Compressed Sensing Measurements," submitted.
- [56] B. Han, F. Wu, and D. Wu, "Image Representation by Compressive Sensing for Visual Sensor Networks," *Journal of Visual Communication and Image Representation*, vol. 21, no. 4, May 2010, pp. 325–333.
- [57] J. Haupt and R. Nowak, "Signal Reconstruction from Noisy Random Projections," *IEEE Transactions on Information Theory*, vol. 52, no. 9, September 2006, pp. 4036–4048.
- [58] L. He and L. Carin, "Exploiting Structure in Wavelet-Based Bayesian Compressive Sensing," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, September 2009, pp. 3488–3497.

- [59] L. He, H. Chen, and L. Carin, “Tree-Structured Compressive Sensing With Variational Bayesian Analysis,” *IEEE Signal Processing Letters*, vol. 17, no. 3, March 2010, pp. 233–236.
- [60] L. J. Hornbeck, “Digital Light Processing for high-brightness high-resolution applications,” 1997, pp. 27–40.
- [61] *Information Technology—Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s*, ISO/IEC 11172-2, 1993, MPEG-1 Video Coding Standard.
- [62] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video*, ISO/IEC 13818-2, 1995, MPEG-2 Video Coding Standard.
- [63] *Information Technology—JPEG 2000 Image Coding System—Part 1: Core Coding System*, ISO/IEC 15444-1, 2000.
- [64] *Advanced Video Coding for Generic Audiovisual Services*, ITU-T, May 2003, ITU-T Recommendation H.264.
- [65] L. Jacques, D. K. Hammond, and J. M. Fadili, “Dequantizing Compressed Sensing: When Oversampling and Non-Gaussian Constraints Combine,” *IEEE Transactions on Information Theory*, vol. 57, no. 1, January 2011, pp. 559–571.
- [66] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici, “CMOS Compressed Imaging by Random Convolution,” *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, April 2009, pp. 1113–1116.
- [67] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [68] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, Englewood Cliffs, NJ, 1984.
- [69] S. Ji, Y. Xue, and L. Carin, “Bayesian Compressive Sensing,” *IEEE Transactions on Signal Processing*, vol. 56, no. 6, June 2008, pp. 2346–2356.
- [70] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, “k-t FOCUSS: A General Compressed Sensing Framework for High Resolution Dynamic MRI,” *Magnetic Resonance in Medicine*, vol. 61, no. 1, January 2009, pp. 103–116.
- [71] H. Jung and J. C. Ye, “Motion Estimated and Compensated Compressed Sensing Dynamic Magnetic Resonance Imaging: What We Can Learn from Video Compression Techniques,” *Imaging Systems and Technology*, vol. 20, no. 2, June 2010, pp. 81–98.



- [72] Y. Kim, M. S. Nadar, and A. Bilgin, “Compressed Sensing Using a Gaussian Scale Mixtures Model in Wavelet Domain,” *Proceedings of the International Conference on Image Processing*, Hong Kong, September 2010, pp. 3365–3368.
- [73] N. G. Kingsbury, “Complex Wavelets for Shift Invariant Analysis and Filtering of Signals,” *Journal of Applied Computational Harmonic Analysis*, vol. 10, May 2001, pp. 234–253.
- [74] L. Landweber, “An Iteration Formula for Fredholm Integral Equations of the First Kind,” *American Journal of Mathematics*, vol. 73, no. 3, July 1951, pp. 615–624.
- [75] H. D. Lüke, “The Origins of the Sampling Theorem,” *IEEE Communications Magazine*, vol. 37, no. 4, April 1999, pp. 106–108.
- [76] M. Lustig, D. Donoho, and J. M. Pauly, “Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging,” *Magnetic Resonance in Medicine*, vol. 58, no. 6, December 2007, pp. 1182–1195.
- [77] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, “Compressed Sensing MRI,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, March 2008, pp. 72–82.
- [78] J. Makhoul, “A Fast Cosine Transform in One and Two Dimensions,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 1, February 1980, pp. 27–34.
- [79] A. Maleki and D. L. Donoho, “Optimally Tuned Iterative Reconstruction Algorithms for Compressed Sensing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, April 2010, pp. 330–341.
- [80] S. Mallat and Z. Zhang, “Matching Pursuits with Time-Frequency Dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, December 1993, pp. 3397–3415.
- [81] R. Marcia, Z. Harmany, and R. Willett, “Compressive Coded Aperture Imaging,” *Computational Imaging III*, San Jose, CA, January 2009, Proc. SPIE 7246, p. 72460G.
- [82] S. Mun and J. E. Fowler, “Block Compressed Sensing of Images Using Directional Transforms,” *Proceedings of the International Conference on Image Processing*, Cairo, Egypt, November 2009, pp. 3021–3024.
- [83] S. Mun and J. E. Fowler, “Residual Reconstruction for Block-Based Compressed Sensing of Video,” *Proceedings of the Data Compression Conference*, J. A. Storer and M. W. Marcellin, eds., Snowbird, UT, March 2011, pp. 183–192.

- [84] S. Mun and J. E. Fowler, “DPCM for Quantized Block-Based Compressed Sensing of Images,” *Proceedings of the European Signal Processing Conference*, Bucharest, Romania, August 2012, to appear.
- [85] H. G. Musmann, “Predictive Image Coding,” *Image Transmission Techniques*, W. K. Pratt, ed., Academic Press, 1979, pp. 73–112.
- [86] D. Needell and J. A. Tropp, “CoSAMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, May 2009, pp. 301–321.
- [87] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, “Image Denoising Using Scale Mixtures of Gaussians in the Wavelet Domain,” *IEEE Transactions on Image Processing*, vol. 12, no. 11, November 2003, pp. 1338–1351.
- [88] R. Robucci, J. D. Gray, L. K. Chiu, J. Romberg, and P. Hasler, “Compressive Sensing on a CMOS Separable-Transform Image Sensor,” *Proceedings of the IEEE*, vol. 98, no. 6, June 2010, pp. 1089–1101.
- [89] J. Romberg, “Compressive Sensing by Random Convolution,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 4, December 2009, pp. 1098–1128.
- [90] P. Schniter, L. C. Potter, and J. Ziniel, “Fast Bayesian Matching Pursuit: Model Uncertainty and Parameter Estimation for Sparse Linear Models,” *IEEE Transactions on Signal Processing*, 2008, submitted.
- [91] C. E. Shannon, “A Mathematical Theory of Communication,” *The Bell System Technical Journal*, vol. 27, July, October 1948, pp. 379–423, 623–656.
- [92] C. E. Shannon, “Communication in the Presence of Noise,” *Proceedings of the IRE*, vol. 37, no. 1, January 1949, pp. 10–21.
- [93] C. E. Shannon, “Coding Theorems for a Discrete Source with a Fidelity Criterion,” *IRE International Convention Record, Part 4*, 1959, vol. 7, pp. 142–163.
- [94] J. M. Shapiro, “Embedded Image Coding Using Zerotrees of Wavelet Coefficients,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, December 1993, pp. 3445–3462.
- [95] G. J. Sullivan, “Multi-hypothesis Motion Compensation for Low Bit-rate Video Coding,” *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, MN, April 1993, vol. 5, pp. 437–440.
- [96] J. Z. Sun and V. K. Goyal, “Optimal Quantization of Random Measurements In Compressed Sensing,” *Proceedings of the IEEE International Symposium on Information Theory*, Seoul, Korea, June 2009, pp. 6–10.

- [97] W. Sweldens, “The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions,” *Wavelet Applications in Signal and Image Processing III*, A. F. Laine, M. A. Unser, and M. V. Wickerhauser, eds., San Diego, CA, July 1995, Proc. SPIE 2569, pp. 68–79.
- [98] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, “A New Compressive Imaging Camera Architecture Using Optical-Domain Compression,” *Computational Imaging IV*, C. A. Bouman, E. L. Miller, and I. Pollak, eds., San Jose, CA, January 2006, Proc. SPIE 6065, p. 606509.
- [99] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, 1996, pp. 267–288.
- [100] M. Trocan, B. Pesquet-Popescu, and J. E. Fowler, “Graph-Cut Rate Distortion Algorithm for Contourlet-Based Image Compression,” *Proceedings of the International Conference on Image Processing*, San Antonio, TX, September 2007, vol. 3, pp. 169–172.
- [101] J. Tropp and A. Gilbert, “Signal Recovery from Random Measurements Via Orthogonal Matching Pursuit,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, December 2007, pp. 4655–4666.
- [102] J. A. Tropp and S. J. Wright, “Computational Methods for Sparse Solution of Linear Inverse Problems,” *Proceedings of the IEEE*, vol. 98, no. 6, June 2010, pp. 948–958.
- [103] Y. Tsaig and D. L. Donoho, “Extensions of Compressed Sensing,” *Signal Processing*, vol. 86, no. 3, March 2006, pp. 549–571.
- [104] N. Vaswani and W. Lu, “Modified-CS: Modifying Compressive Sensing for Problems with Partially Known Support,” *IEEE Transactions on Signal Processing*, vol. 58, no. 9, September 2010, pp. 4595–4607.
- [105] R. G. Vaughan, N. L. Scott, and D. R. White, “The Theory of Bandpass Sampling,” *IEEE Transactions on Signal Processing*, vol. 39, no. 9, September 1991, pp. 1973–1984.
- [106] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, “An Architecture for Compressive Imaging,” *Proceedings of the International Conference on Image Processing*, Atlanta, GA, October 2006, pp. 1273–1276.
- [107] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, “Compressive Imaging for Video Representation and Coding,” *Proceedings of the Picture Coding Symposium*, Beijing, China, April 2006.

- [108] L. Wang, X. Wu, and G. Shi, “Progressive Quantization of Compressive Sensing Measurements,” *Proceedings of the Data Compression Conference*, J. A. Storer and M. W. Marcellin, eds., Snowbird, UT, March 2011, pp. 233–242.
- [109] L. Wang, X. Wu, and G. Shi, “Binned Progressive Quantization for Compressive Sensing,” *IEEE Transactions on Image Processing*, vol. 21, no. 6, June 2012, pp. 2980–2990.
- [110] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, “Sparse Reconstruction by Separable Approximation,” *IEEE Transactions on Signal Processing*, vol. 57, no. 7, July 2009, pp. 2479–2493.
- [111] X. Wu, X. Zhang, and J. Wang, “Model-Guided Adaptive Recovery of Compressive Sensing,” *Proceedings of the Data Compression Conference*, J. A. Storer and M. W. Marcellin, eds., Snowbird, UT, March 2009, pp. 123–132.