Block-Based Compressed Sensing of Images and Video
Block-Based Compressed Sensing of Images and Video

James E. Fowler
Mississippi State University
USA
fowler@ece.msstate.edu

Sungkwang Mun
Mississippi State University
USA
sm655@msstate.edu

Eric W. Tramel
Mississippi State University
USA
ewt16@msstate.edu

Boston – Delft
Foundations and Trends® in
Signal Processing
Volume 4 Issue 4, 2010
Editorial Board

Editor-in-Chief:
Robert M. Gray
Dept of Electrical Engineering
Stanford University
350 Serra Mall
Stanford, CA 94305
USA
rmgray@stanford.edu

Editors
Abeer Alwan (UCLA) Jelena Kovacevic (CMU)
John Apostolopoulos (HP Labs) Jia Li (Pennsylvania State University)
Pamela Cosman (UCSD) B.S. Manjunath (UCSB)
Michelle Effros (California Institute of Technology) Urbashi Mitra (USC)
Yonina Eldar (Technion) Thrasos Pappas (Northwestern University)
Yariv Ephraim (George Mason University) Mihaela van der Shaar (UCLA)
Sadaoki Furui (Tokyo Institute of Technology) Michael Unser (EPFL)
Vivek Goyal (MIT) P.P. Vaidyanathan (California Institute of Technology)
Sinan Gunturk (Courant Institute) Rabab Ward (University of British Columbia)
Christine Guillemot (IRISA) Susie Wee (HP Labs)
Shella Hemami (Cornell) Clifford J. Weinstein (MIT Lincoln Laboratories)
Lina Karam (Arizona State University) Min Wu (University of Maryland)
Nick Kingsbury (Cambridge University) Josiane Zerubia (INRIA)
Alex Kot (Nanyang Technical University) Pao-Chi CHang (National Central University)

Full text available at: http://dx.doi.org/10.1561/2000000033
Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
- classification and detection
- estimation and regression
- tree-structured methods

Information for Librarians
Foundations and Trends® in Signal Processing, 2010, Volume 4, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.
Block-Based Compressed Sensing of Images and Video

James E. Fowler¹, Sungkwang Mun² and Eric W. Tramel³

¹ Department of Electrical & Computer Engineering, Geosystems Research Institute, Mississippi State University, Mississippi State, MS 39762, USA, fowler@ece.msstate.edu
² Department of Electrical & Computer Engineering, Geosystems Research Institute, Mississippi State University, Mississippi State, MS 39762, USA, sm655@msstate.edu
³ Department of Electrical & Computer Engineering, Geosystems Research Institute, Mississippi State University, Mississippi State, MS 39762, USA, ewt16@msstate.edu

Abstract

A number of techniques for the compressed sensing of imagery are surveyed. Various imaging media are considered, including still images, motion video, as well as multiview image sets and multiview video. A particular emphasis is placed on block-based compressed sensing due to its advantages in terms of both lightweight reconstruction complexity as well as a reduced memory burden for the random-projection measurement operator. For multiple-image scenarios, including video and multiview imagery, motion and disparity compensation is employed to exploit frame-to-frame redundancies due to object motion and parallax, resulting in residual frames which are more compressible
and thus more easily reconstructed from compressed-sensing measurements. Extensive experimental comparisons evaluate various prominent reconstruction algorithms for still-image, motion-video, and multiview scenarios in terms of both reconstruction quality as well as computational complexity.
# Contents

**Acronyms**

1 Introduction 5

2 Compressed Sensing 9

2.1 An Overview of CS Theory 9

2.2 Approaches to CS-Based Signal Acquisition 11

2.3 Approaches to CS Reconstruction 13

2.4 CS versus Source Coding 15

3 Block-Based Compressed Sensing for Still Images 17

3.1 CS Acquisition of Still Images 17

3.2 Straightforward Reconstruction for Images 22

3.3 Total-Variation Reconstruction 24

3.4 CS with Blocks in the Spatial Domain 25

3.5 CS with Blocks in the Wavelet-Domain 35

3.6 Other Approaches to CS Reconstruction of Images 39

3.7 Comparison of Various CS Techniques for Images 41

3.8 Perspectives 45

4 Block-Based Compressed Sensing for Video 47

4.1 CS Acquisition of Video 48

Full text available at: http://dx.doi.org/10.1561/2000000033
5 Multihypothesis Prediction for Compressed Sensing of Video

5.1 Prediction Strategies for Residual Reconstruction 78
5.2 SH Frame Prediction for CS Reconstruction 80
5.3 MH Frame Prediction for CS Reconstruction 82
5.4 An Alternate $\ell_1$-Based MH Regularization 84
5.5 Experimental Observations 85
5.6 Comparison of Various CS Techniques for Video 89
5.7 Perspectives 93

6 Compressed Sensing of Multiview Image and Video

6.1 Single-View Reconstruction 98
6.2 Multistage Reconstruction of Multiview Images 100
6.3 Reconstruction of Multiview Video 101
6.4 Experimental Observations 103
6.5 Perspectives 106

7 Conclusions 109

Acknowledgments 111

References 113
Acronyms

3D-BCS-SPL three-dimensional BCS-SPL. 50, 58, 63, 65, 68

BCS block-based compressed sensing. 3, 21–24, 27–32, 34, 37, 38, 41, 42, 46, 48, 51, 77, 81, 99

BCS-GPSR-DCT block-based compressed sensing with GPSR reconstruction using a block-based DCT sparsity basis. 27, 31

BCS-GPSR-DWT block-based compressed sensing with GPSR reconstruction using a whole-image DWT sparsity basis. 27, 31

BCS-SPL block-based compressed sensing with smooth projected Landweber reconstruction. 24, 32, 34, 35, 37, 39, 41, 42, 49–53, 55, 58, 60, 63, 65, 68, 71, 74, 81, 82, 84, 85, 87, 92, 94–96, 98, 99

BCS-SPL-CT block-based compressed sensing with smooth projected Landweber reconstruction using a CT sparsity basis. 30, 31

BCS-SPL-DCT block-based compressed sensing with smooth projected Landweber reconstruction using a block-based DCT sparsity basis. 29, 31
2. **Acronyms**

- **BCS-SPL-DDWT** block-based compressed sensing with smooth projected Landweber reconstruction using a DDWT sparsity basis. [30, 31]
- **BCS-SPL-DWT** block-based compressed sensing with smooth projected Landweber reconstruction using a whole-image DWT sparsity basis. [29, 31]
- **BCS-TV** block-based compressed sensing with total-variation reconstruction. [29, 30–32]
- **BP** basis pursuit. [9, 10]
- **BPDN** basis-pursuit denoising. [10, 20]
- **CoSaMP** compressive sampling matching pursuit. [10, 36]
- **CT** contourlet transform. [25, 26, 29]
- **DC** disparity compensation. [94, 98, 101–103]
- **DC-BCS-SPL** disparity-compensated BCS-SPL. [94, 99, 102]
- **DCT** discrete cosine transform. [19, 27, 29, 38, 49, 63, 64]
- **DDWT** dual-tree discrete wavelet transform. [26, 29, 38, 85, 99]
- **DE** disparity estimation. [94, 99, 101, 103]
- **DISCOS** distributed compressed video sensing. [61, 75, 80]
- **DLP** digital-light-processing. [14]
- **DMD** digital micromirror device. [14, 15, 22, 45, 48]
- **DWT** discrete wavelet transform. [19, 23, 25, 27, 29, 32, 34, 36, 38, 62, 85, 99]
- **GOP** group of pictures. [49, 50, 54, 58, 60, 66, 69, 85, 92]
- **GPSR** gradient projection for sparse reconstruction. [10, 27, 29, 36, 38, 60, 82]
- **IST** iterative splitting and thresholding. [10]
- **JL** Johnson-Lindenstrauss. [76]
- **k–t FOCUSS** focal underdetermined system solver in k–t space. [61, 62, 71, 73, 76, 85, 86, 90, 92]
**Acronyms**

LASSO least absolute shrinkage and selection operator. 9
LDS linear dynamical system. 59
LIMAT lifting-based invertible motion adaptive transform. 62
LSQ least-squares. 78, 79

MARX model-based adaptive recovery of compressive sensing. 36, 37
MC motion compensation. 44, 50, 52, 54, 58, 63, 66, 69, 70, 73, 75, 76
MC-BCS-SPL motion-compensated BCS-SPL. 50, 52, 71, 75, 76, 85
MCTF motion-compensated temporal filtering. 62, 63
ME motion estimation. 44, 50, 52, 58, 63, 66, 69, 70, 73, 75, 76, 81, 83
MEMS microelectromechanical systems. 45
MH multihypothesis. 73, 78, 85, 90, 92
MH-BCS-SPL multihypothesis BCS-SPL. 83, 88, 92
MRI magnetic resonance imaging. 2, 58, 62, 86, 89, 92
MS multiscale. 32, 86
MS-BCS-SPL multiscale, wavelet-domain block-based compressed sensing with smooth projected Landweber reconstruction. 33–35, 37, 39, 41, 87, 89, 91
MS-GPSR multiscale GPSR. 36, 38, 41

OMP orthogonal matching pursuits. 10, 36

PAR piecewise autoregressive. 36
PL projected Landweber. 11, 24, 29
PSNR peak signal-to-noise ratio. 27, 31, 38, 39, 64, 69, 82, 84, 86, 90, 92, 100, 101

SALSA split augmented Lagrangian shrinkage algorithm. 37, 41
SAMP sparsity adaptive matching pursuits. 36
SH single-hypothesis. 73, 77, 82, 84
SpaRSA sparse reconstruction via separable approximation. 10
SPL smoothed projected Landweber. 24, 26
4 Acronyms

SRM structurally random matrix. 8, 15, 18, 20, 21, 32, 37, 38, 40, 42, 87

StOMP stagewise orthogonal matching pursuit. 36

TSW-CS tree-structured wavelet compressed sensing. 36, 37, 41

TV total variation. 21, 23, 28, 30, 31, 37, 42, 87, 91
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 [109], the sampling theorem can be traced to earlier work by telegraphers and mathematicians (see, e.g., [82]). 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 [8]. 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., [129]) 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)\(^1\) (e.g., [18, 22, 38]) 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 nonlinear 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.

The goal of this monograph is to overview some of these methods. A primary focus is an examination of the state of the art in CS reconstruction for various imaging modalities, including still images, motion video, and multiview imagery. Throughout, 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., [83, 84]). On the other hand, CS for photographic imagery

\(^1\)Also known as compressive sampling or compressive sensing.
acquired in the spatial domain is a comparatively emerging area and, thus, the topic of the present monograph.

An additional focus of this monograph is on CS reconstruction as applied on image blocks. 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. These 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 note also that our discussion is not intended to serve as an in-depth tutorial on the theory or mathematics of CS; rather, there exist several excellent overviews on this subject (e.g., [7, 20, 22]). Instead, our coverage of CS theory here will be brief, while the specifics of the application of BCS to natural imagery will consume the bulk of the discussion.

The remainder of the monograph is organized as follows. Section 2 briefly overviews CS theory, including acquisition and reconstruction processes. Section 3 then considers the CS of a single still image, focusing on a variety of techniques to reconstruct such images from random CS measurements. Section 4 extends these concepts to the CS of video with an emphasis on reconstruction from motion-compensated residuals, and then Section 5 adds multihypothesis prediction to such motion-based CS reconstruction. Section 6 finally considers the CS of multiview images and video in which a scene is imaged from several viewpoints simultaneously. We end the monograph by making several concluding remarks.
References


References


References


References


References


[67] 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, MPEG-1 Video Coding Standard, 1993.


References


Full text available at: http://dx.doi.org/10.1561/2000000033
References


References


Full text available at: http://dx.doi.org/10.1561/2000000033
References


