BLOCK COMPRESSED SENSING OF IMAGES USING DIRECTIONAL TRANSFORMS

Sungkwang Mun and James E. Fowler
Department of Electrical and Computer Engineering, Geosystems Research Institute,
Mississippi State University, USA

ABSTRACT

Block-based random image sampling is coupled with a projection-driven compressed-sensing recovery that encourages sparsity in the domain of directional transforms simultaneously with a smooth reconstructed image. Both contourlets as well as complex-valued dual-tree wavelets are considered for their highly directional representation, while bivariate shrinkage is adapted to their multiscale decomposition structure to provide the requisite sparsity constraint. Smoothing is achieved via a Wiener filter incorporated into iterative projected Landweber compressed-sensing recovery, yielding fast reconstruction. The proposed approach yields images with quality that matches or exceeds that produced by a popular, yet computationally expensive, technique which minimizes total variation. Additionally, reconstruction quality is substantially superior to that from several prominent pursuits-based algorithms that do not include any smoothing.

Index Terms— Compressed sensing, contourlets, dual-tree discrete wavelet transform, bivariate shrinkage

1. INTRODUCTION

Recent years have seen significant interest in the paradigm of compressed sensing (CS) [1–3] which permits, under certain conditions, signals to be sampled at sub-Nyquist rates via linear projection onto a random basis while still enabling exact reconstruction of the original signal. As applied to 2D images, however, CS faces several challenges including a computationally expensive reconstruction process and huge memory required to store the random sampling operator. Recently, several fast algorithms (e.g., [4–6]) have been developed for CS reconstruction, while the latter challenge was addressed in [7] using a block-based sampling operation. Additionally in [7], projection-based Landweber iterations were proposed to accomplish fast CS reconstruction while simultaneously imposing smoothing with the goal of improving the reconstructed-image quality by eliminating blocking artifacts.

In this paper, we adopt this same basic framework of block-based CS sampling of images coupled with iterative projection-based reconstruction with smoothing. Our contribution lies in that we cast the reconstruction in the domain of recent transforms that feature a highly directional decomposition. These transforms—specifically, contourlets [8] and complex-valued dual-tree wavelets [9]—have shown promise to overcome deficiencies of widely-used wavelet transforms in several application areas. In their application to iterative projection-based CS recovery, we adapt bivariate shrinkage [10] to their directional decomposition structure to provide sparsity-enforcing thresholding, while a Wiener-filter step encourages smoothness of the result. In experimental simulations, we find that the proposed CS reconstruction based on directional transforms outperforms equivalent reconstruction using common wavelet and cosine transforms. Additionally, the proposed technique usually matches or exceeds the quality of total-variation (TV) reconstruction [11], a popular approach to CS recovery for images whose gradient-based operation also promotes smoothing but runs several orders of magnitude slower than our proposed algorithm.

2. BACKGROUND

Suppose that we want to recover real-valued signal $\mathbf{x}$ with length N from M samples, $M \ll N$; i.e., we want to recover $\mathbf{x}$ from $\mathbf{y} = \mathbf{\Phi x}$, where $\mathbf{y}$ has length M, and $\mathbf{\Phi}$ is an $M \times N$ measurement matrix. Because the number of unknowns is much larger than observations, recovering every $\mathbf{x} \in \Re^N$ from its corresponding $\mathbf{y}$ is impossible in general; however, if $\mathbf{x}$ is sufficiently *sparse*, exact recovery is possible—this is the fundamental tenant of CS theory; see, e.g., [3], for a more complete overview. The usual choice for the measurement basis $\mathbf{\Phi}$ is a random matrix; here, we further assume that $\mathbf{\Phi}$ is orthonormal such that $\mathbf{\Phi\Phi}^T = \mathbf{I}$.

Quite often, the requisite sparsity will exist with respect to some transform $\mathbf{\Psi}$. In this case, the key to CS recovery is the production of a sparse set of significant transform coefficients, $\check{\mathbf{x}} = \mathbf{\Psi x}$, and the ideal recovery procedure searches for the $\check{\mathbf{x}}$ with the smallest ℓ_0 norm consistent with the observed $\mathbf{y}$. However, this ℓ_0 optimization being NP-complete, several alternative solution procedures have been proposed. Perhaps the most prominent of these is basis pursuit (BP) [12] which applies a convex relaxation to the ℓ_0 problem resulting in an ℓ_1 optimization,

$$\check{\mathbf{x}} = \arg\min_{\check{\mathbf{x}}} \|\check{\mathbf{x}}\|_1, \quad \text{such that } \mathbf{y} = \mathbf{\Phi\Psi}^{-1}\check{\mathbf{x}}, \tag{1}$$

where $\mathbf{\Psi}^{-1}$ is the inverse transform. Although BP can be implemented effectively with linear programming, its computational complexity is often high, leading to recent interest in reduced-complexity relaxations (e.g., gradient projection for sparse reconstruction (GPSR) [4]) as well as in greedy BP variants, including matching pursuits, orthogonal matching pursuits, and, recently, sparsity adaptive matching pursuits (SAMP) [5]. Such algorithms significantly reduce computational complexity at the cost of lower reconstruction quality.

As an alternative to the pursuits class of CS reconstruction, techniques based on projections have been proposed recently (e.g., [6]). Algorithms of this class form $\check{\mathbf{x}}$ by successively projecting and thresholding; for example, the reconstruction in [6] starts from some initial approximation $\check{\mathbf{x}}^{(0)}$ and forms the approximation at iteration $i+1$ as

$$\check{\check{\mathbf{x}}}^{(i)} = \check{\mathbf{x}}^{(i)} + \frac{1}{\gamma}\mathbf{\Psi\Phi}^T\left(\mathbf{y} - \mathbf{\Phi\Psi}^{-1}\check{\mathbf{x}}^{(i)}\right), \tag{2}$$

$$\check{\mathbf{x}}^{(i+1)} = \begin{cases} \check{\check{\mathbf{x}}}^{(i)}, & \left|\check{\check{\mathbf{x}}}^{(i)}\right| \geq \tau^{(i)}, \\ 0 & \text{else.} \end{cases} \tag{3}$$

In *Proceedings of the International Conference on Image Processing*, Cairo, Egypt, November 2009, pp. 3021-3024.

Here, γ is a scaling factor ([6] uses the largest eigenvalue of $\mathbf{\Phi}^T\mathbf{\Phi}$), while $\tau^{(i)}$ is a threshold set appropriately at each iteration. It is straightforward to see that this procedure is a specific instance of a projected Landweber (PL) algorithm [13]. Like the greedy algorithms of the pursuits class, PL-based CS reconstruction also provides reduced computational complexity. Additionally, and perhaps more importantly, the PL formulation offers the possibility of easily incorporating additional optimization criteria. For example, the technique that we overview in the next section incorporates Wiener filtering into the PL iteration to search for a CS reconstruction simultaneously achieving sparsity and smoothness.

3. BLOCK-BASED CS WITH SMOOTHED PL RECONSTRUCTION

In [7], a paradigm for the CS of 2D images was proposed. In this technique, the sampling 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)–(3) that incorporates a smoothing operation. Due to its combination of block-based CS (BCS) sampling and smoothed-PL (SPL) reconstruction, we refer to the overall technique as BCS-SPL. We overview its constituent components below.

3.1. BCS—Block-Based CS Sampling

In BCS, an image is divided into $B \times B$ blocks and sampled 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 $\mathbf{y}_j = \mathbf{\Phi}_B \mathbf{x}_j$, where $\mathbf{\Phi}_B$ is an $M_B \times B^2$ orthonormal measurement matrix with $M_B = \lfloor \frac{M}{N} B^2 \rfloor$.

Using BCS rather than random sampling applied to the entire image $\mathbf{x}$ has several merits [7]. First, the measurement operator $\mathbf{\Phi}_B$ is conveniently stored and employed because of its compact size. Second, the encoder does not need to wait until the entire image is measured, but may send each block after its linear projection. Last, an initial approximation $\mathbf{x}^{(0)}$ with minimum mean squared error can be feasibly calculated due to the small size of $\mathbf{\Phi}_B$ [7]. As done in [7], we employ blocks of size $B = 32$.

3.2. SPL—A Smoothed PL Variant

In [7], Wiener filtering was incorporated into the basic PL framework described in Sec. 2 in order to remove blocking artifacts. In essence, this operation imposes smoothness in addition to the sparsity inherent to PL. Specifically, in [7], a Wiener-filtering step was interleaved with the PL projection of (2)–(3); thus, the approximation to the image at iteration $i+1$, $\mathbf{x}^{(i+1)}$, is produced from $\mathbf{x}^{(i)}$ as:

```
function x^(i+1) = SPL(x^(i), y, Φ_B, Ψ, λ)
  x̂^(i) = Wiener(x^(i))
  for each block j
    x̂̂_j^(i) = x̂_j^(i) + Φ_B^T (y − Φ_B x̂_j^(i))
  x̌̌^(i) = Ψ x̂̂^(i)
  x̌^(i) = Threshold(x̌̌^(i), λ)
  x̄^(i) = Ψ^(−1) x̌^(i)
  for each block j
    x_j^(i+1) = x̄_j^(i) + Φ_B^T (y − Φ_B x̄_j^(i))
```

Here, $\mathrm{Wiener}(\cdot)$ is pixelwise adaptive Wiener filtering using a neighborhood of 3×3, while $\mathrm{Threshold}(\cdot)$ is a thresholding process as discussed below. In our use of SPL, we initialize with $\mathbf{x}^{(0)} = \mathbf{\Phi}^T \mathbf{y}$ and terminate when $|D^{(i+1)} - D^{(i)}| < 10^{-4}$, where $D^{(i)} = \frac{1}{\sqrt{N}} \|\mathbf{x}^{(i)} - \hat{\hat{\mathbf{x}}}^{(i-1)}\|_2$.

4. DIRECTIONAL TRANSFORMS AND BCS-SPL

4.1. Transforms

In [7], several iterations of the $\mathrm{SPL}(\cdot)$ procedure described above are used as an initial step in a dual-stage algorithm for CS reconstruction. The stages employ PL iterations in the form of (2)–(3) using several different transforms, including a block-based lapped cosine transform as well as a redundant wavelet transform. For reasons of simplicity, we now depart from this methodology and instead focus on a single stage of $\mathrm{SPL}(\cdot)$ iterations. This allows us to incorporate several prominent directional transforms into the basic SPL formulation to evaluate their relative efficacy at CS reconstruction. Although we do not pursue it here, multiple SPL stages in the style of [7] could be employed along with these directional transforms to potentially refine performance.

Although the discrete wavelet transform (DWT) is widely used for image compression, 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) [8] preserves interesting features of the traditional DWT, namely multiresolution and local characteristics of the signal, and, at the expense of a 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 DWT (DDWT) [9] 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. DDWT yields a decomposition with a much higher degree of directionality than that possessed by the traditional DWT; 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 for image compression (e.g., [14–16]). The experimental results below explore the efficacy of these directional transforms in the SPL-based CS reconstruction of Sec. 3.

4.2. Thresholding

As originally described in [7], $\mathrm{SPL}(\cdot)$ used hard thresholding in the form of (3). To set a proper τ for hard thresholding, we employ the universal threshold method of [17]. Specifically, in (3),

$$\tau^{(i)} = \lambda \sigma^{(i)} \sqrt{2 \log K}, \tag{4}$$

where λ is a constant control factor to manage convergence, and K is the number of the transform coefficients. As in [17], $\sigma^{(i)}$ is estimated using a robust median estimator,

$$\sigma^{(i)} = \frac{\mathrm{median}(|\check{\check{\mathbf{x}}}^{(i)}|)}{0.6745}. \tag{5}$$

In *Proceedings of the International Conference on Image Processing*, Cairo, Egypt, November 2009, pp. 3021-3024.

Hard thresholding inherently assumes independence between coefficients. However, bivariate shrinkage [10] 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 [10], 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, given a specific transform coefficient ξ and its parent coefficient ξ_p in the next coarser scale, the $\text{Threshold}(\cdot)$ operator in SPL is the MAP estimator of ξ,

$$\text{Threshold}(\xi, \lambda) = \frac{\left(\sqrt{\xi^2 + \xi_p^2} - \lambda \frac{\sqrt{3\sigma^{(i)}}}{\sigma_\xi}\right)_+}{\sqrt{\xi^2 + \xi_p^2}} \cdot \xi, \quad (6)$$

where $(g)_+ = 0$ for $g < 0$, $(g)_+ = g$ else; $\sigma^{(i)}$ is the median estimator of (5) applied to only the finest-scale transform coefficients; and, again, λ is a convergence-control factor. Here, σ_ξ^2 is the marginal variance of coefficient ξ estimated in a local 3×3 neighborhood surrounding ξ as in [10].

5. EXPERIMENTAL RESULTS

To evaluate directional transforms for CS reconstruction, we deploy both the CT and DDWT within the BCS-SPL framework described in Sec. 3. We refer to the resulting implementations as BCS-SPL-CT and BCS-SPL-DDWT, respectively. To evaluate the effectiveness of the increased directionality of the CT and DDWT, we compare to BCS-SPL-DWT, an equivalent approach using the ubiquitous biorthogonal 9-7 DWT. We also compare to a BCS-SPL variant using a block-based DCT for SPL reconstruction; the resulting algorithm (BCS-SPL-DCT) is similar to that initially proposed as the first-stage reconstruction in [7], although a lapped transform was used there. In SPL, we use bivariate shrinkage (6) 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 (4) with $\lambda = 6$.

We compare also to BCS-TV wherein the block-based BCS is still used for image sampling while the SPL reconstruction is replaced with the minimum TV optimization of [11]. Like SPL, such TV-based reconstruction also imposes sparsity and smoothness constraints; unlike the explicit smoothing of SPL's Wiener filtering, however, smoothness in TV is implicit in that the solution is sparse in a gradient space. We use ℓ_1-MAGIC[1] in the BCS-TV implementation. Finally, as representative of the pursuits class of CS reconstruction, we compare to GPSR[2] [4] as well as SAMP[3] [5]; for these, we use implementations provided by their respective authors.

Table 1 compares PSNR for several 512×512 images at several measurement ratios, M/N. We note that, since the quality of reconstruction can vary due to the randomness of the measurement matrix $\mathbf{\Phi}$ (or $\mathbf{\Phi}_B$), 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 measurement 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. However, the ℓ_1-MAGIC implementation of TV is quite slow—BCS-TV takes 3–4 hours for each trial, whereas the BCS-SPL implementations run for only 1–5 minutes depending on the complexity of the transform used. These later times are in line with GPSR (less than 60 seconds) and SAMP (several minutes). Times are for a 3.2-GHz dual-core processor.

Fig. 1 illustrates example visual results. We note that, despite the smoothing inherent to TV reconstruction, blocking artifacts are apparent. On the other hand, the smoothing of the BCS-SPL-DDWT reconstruction (BCS-SPL-CT yields similar visual quality) eliminates blocking while the enhanced directionality of the transform provides better quality than the DWT- and DCT-based techniques. Finally, we note that the pursuits-based algorithms which do not factor in any sort of smoothing—GPSR and SAMP—yield images of noticeably deficient visual quality.

[1] `http://www.acm.caltech.edu/l1magic/`

[2] `http://www.lx.it.pt/~mtf/GPSR/`

[3] `http://thongdojhu.googlepages.com/samp_intro/`

6. CONCLUSIONS

In this paper, we examined the use of recently proposed directional transforms in the CS reconstruction of images. We adopted the general paradigm of block-based random image sampling coupled with a projection-based reconstruction promoting not only sparsity but also smoothness of the reconstruction. This framework facilitates the incorporation into the CS-recovery process of directional transforms based on contourlets and complex-valued dual-tree wavelets. The resulting algorithms inherit the fast execution speed of the projection-based CS reconstruction while the enhanced directionality coupled with a smoothing step encourages superior image quality, particularly at low sampling rates.

7. REFERENCES

[1] E. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, December 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

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

[4] 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, pp. 586–597, December 2007.

[5] T. T. Do, L. Gan, N. Nguyen, and T. D. Tran, "Sparsity adaptive matching pursuit algorithm for practical compressed sensing," in *Proceedings of the* 42^{th} *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, October 2008, pp. 581–587.

[6] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4036–4048, September 2006.

Figure 1: Lenna for $M/N = 20\%$. Left to right: BCS-SPL-DDWT, PSNR = 31.37 dB; BCS-SPL-DCT, PSNR = 30.45 dB; BCS-TV, PSNR = 30.60 dB; SAMP, PSNR = 28.54 dB.

Table 1: PSNR Performance in dB

Algorithm	*Measurement Rate (M/N)* 0.1	0.2	0.3	0.4	0.5
Lenna					
BCS-SPL-CT	28.17	31.02	32.99	34.68	36.25
BCS-SPL-DDWT	**28.31**	**31.37**	**33.50**	**35.20**	**36.78**
BCS-SPL-DWT	27.81	30.89	32.94	34.61	36.15
BCS-SPL-DCT	27.70	30.45	32.46	34.19	35.77
BCS-TV	27.86	30.60	32.56	34.25	35.89
SAMP	25.94	28.54	32.04	33.93	35.37
GPSR	24.69	28.54	31.53	33.69	35.82
Barbara					
BCS-SPL-CT	22.75	**24.33**	25.90	27.54	29.38
BCS-SPL-DDWT	**22.85**	24.29	**25.92**	27.50	29.12
BCS-SPL-DWT	22.62	23.94	25.20	26.56	28.05
BCS-SPL-DCT	22.76	24.38	25.91	27.42	29.05
BCS-TV	22.45	23.60	24.57	25.57	26.73
SAMP	20.97	22.83	25.04	**27.68**	30.08
GPSR	20.23	22.66	24.99	27.42	**30.15**
Peppers					
BCS-SPL-CT	28.56	31.04	32.57	33.77	34.88
BCS-SPL-DDWT	**28.88**	**31.44**	**32.89**	**34.06**	**35.18**
BCS-SPL-DWT	28.69	31.04	32.48	33.63	34.74
BCS-SPL-DCT	27.88	30.41	31.90	33.08	34.20
BCS-TV	28.52	31.21	32.74	33.96	35.17
SAMP	25.94	28.61	30.69	31.71	32.42
GPSR	24.58	28.19	30.19	31.76	33.21
Mandrill					
BCS-SPL-CT	22.87	**24.97**	**26.95**	**28.90**	**30.93**
BCS-SPL-DDWT	**22.94**	24.87	26.69	28.42	30.28
BCS-SPL-DWT	22.54	24.33	26.03	27.68	29.38
BCS-SPL-DCT	22.31	24.15	25.94	27.77	29.68
BCS-TV	22.31	24.34	26.08	27.77	29.45
SAMP	20.20	21.56	24.00	27.10	30.29
GPSR	20.11	22.23	24.37	26.99	30.06
Goldhill					
BCS-SPL-CT	26.85	**28.95**	30.48	31.92	33.28
BCS-SPL-DDWT	**26.96**	28.93	30.45	31.79	33.11
BCS-SPL-DWT	26.71	28.68	30.13	31.53	32.85
BCS-SPL-DCT	26.10	28.32	29.63	30.98	32.57
BCS-TV	26.53	28.85	**30.56**	**32.09**	**33.61**
SAMP	24.31	26.30	28.07	29.45	30.86
GPSR	23.63	26.14	28.09	30.02	31.72

[7] L. Gan, "Block compressed sensing of natural images," in *Proceedings of the International Conference on Digital Signal Processing*, Cardiff, UK, July 2007, pp. 403–406.

[8] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, December 2005.

[9] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied Computational Harmonic Analysis*, vol. 10, pp. 234–253, May 2001.

[10] 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, pp. 2744–2756, November 2002.

[11] 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, pp. 1207–1223, August 2006.

[12] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, August 1998.

[13] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Bristol, UK: Institute of Physics Publishing, 1998.

[14] M. Trocan, B. Pesquet-Popescu, and J. E. Fowler, "Graph-cut rate distortion algorithm for contourlet-based image compression," in *Proceedings of the International Conference on Image Processing*, vol. 3, San Antonio, TX, September 2007, pp. 169–172.

[15] J. E. Fowler, J. B. Boettcher, and B. Pesquet-Popescu, "Image coding using a complex dual-tree wavelet transform," in *Proceedings of the European Signal Processing Conference*, Poznań, Poland, September 2007.

[16] 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, pp. 633–636, September 2007.

[17] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.