COMPRESSED SENSING OF MULTIVIEW IMAGES USING DISPARITY COMPENSATION

Maria Trocan[†], Thomas Maugey[‡], Eric W. Tramel[*], James E. Fowler[*], Béatrice Pesquet-Popescu[‡]

[†]*Institut Supérieur d'Electronique de Paris*, [‡]*Télécom ParisTech*, [*]*Mississippi State University*

ABSTRACT

Compressed sensing is applied to multiview image sets and inter-image disparity compensation is incorporated into image reconstruction in order to take advantage of the high degree of inter-image correlation common to multiview scenarios. Instead of recovering images in the set independently from one another, two neighboring images are used to calculate a prediction of a target image, and the difference between the original measurements and the compressed-sensing projection of the prediction is then reconstructed as a residual and added back to the prediction in an iterated fashion. The proposed method shows large gains in performance over straightforward, independent compressed-sensing recovery. Additionally, projection and recovery are block-based to significantly reduce computation time.

Index Terms— Compressed sensing, multiview images, disparity compensation

1. INTRODUCTION

Many systems today use multiple cameras to capture information about a specified scene, such as 3D reconstruction, creation of virtual environments, and surveillance applications. Because multiview systems require multiple sensors, the cost of data acquisition is often much higher than that of traditional systems. In these multiple perspective, or multiview, situations, the correlation between images is often very high due to similar content. Compression, restoration, or other data-processing tasks can benefit greatly by exploiting this redundancy of content to improve their performance. Disparity compensation (DC) between the images within a multiview image set can be used to take advantage of this correlation.

Compressed sensing (CS) (e.g. [1]) is a recent paradigm which allows for a signal to be sampled at sub-Nyquist rates and proposes a methodology of recovery which incurs no loss. CS tells us that this is achievable under the assumption that the original signal can be described sparsely in either its ambient domain or in some other basis, Ψ. The core of the signal-acquisition step commonly involves a projection onto a random basis, Φ, which must exhibit a high level of incoherence with the sparse domain [1]. Physical implementations of this methodology have been made, such as the well-known single-pixel camera [2], and many methods have been proposed for the recovery of signals acquired in this manner [3–8].

In this paper, we propose a joint CS reconstruction algorithm for multiview image sets which takes advantage of the strong correlation between images within the set. In [4], an efficient algorithm for reconstructing randomly projected blocked images was proposed. The goal of this paper is to enhance the accuracy of this algorithm within the multiview setting through the use of inter-image DC during the reconstruction process. The results we obtain are promising and show substantial performance improvement over the straightforward, independent CS recovery of the images of the set, even at very low subsampling rates.

2. PRELIMINARIES

One of the main advantages of the CS paradigm is the very low computational burden placed on the encoding process, which requires only the projection of the signal $\mathbf{x}$, of dimensionality N, onto some measurement basis, $\Phi_{N \times M}$, where $M \ll N$. The result of this computation is the M-dimensional vector of measurements, $\mathbf{y} = \Phi \mathbf{x}$. Φ is often chosen to be a random matrix because it satisfies the incoherency requirements of CS reconstruction for any structured signal transform Ψ with a high probability. In this way, the encoder can also be said to be structure agnostic. We assume Φ is also chosen to be orthonormal ($\Phi^T \Phi = \mathbf{I}$).

This light encoding procedure offloads most the computation of CS onto the decoder. Because the inverse of the projection $\hat{\mathbf{x}} = \Phi^{-1} \mathbf{y}$ is ill-posed, we cannot directly solve the inverse problem to find the original signal from the given measurements. Instead, the CS paradigm tells us that the correct solution for $\mathbf{x}$ is the sparsest signal which lies in the set of signals that match the measurements [1]; i.e.,

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \Psi \mathbf{x} \right\|_{\ell_0} \qquad \text{s.t.} \qquad \mathbf{y} = \Phi \mathbf{x}, \qquad (1)$$

where sparsity is measured in the domain of transform Ψ. However, this ℓ_0-constrained optimization problem is computationally infeasible due to its combinational and non-differentiable nature. Thus, a ℓ_1 convex relaxation is often applied, sacrificing accuracy but permitting the recovery to be implemented directly via linear-programming techniques (e.g., [7–9]). Further relaxations of the optimization have also been attempted, such as the mixed ℓ_1-ℓ_2 method proposed in [10]. However, all of these schemes still suffer from very long reconstructions times for N of any practical or interesting size.

Iterative thresholding algorithms have also been proposed as another class of solutions for CS recovery. The most common of these is the iterated hard thresholding (IHT) algorithm (e.g., [11–14]). IHT replaces the constrained optimization formulation with an unconstrained optimization problem via a Lagrangian multiplier and further relaxes the problem by loosening the equality constraint to an ℓ_2-distance penalty,

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\| \Psi \mathbf{x} \right\|_{\ell_1} + \lambda \left\| \mathbf{y} - \Phi \mathbf{x} \right\|_{\ell_2}. \qquad (2)$$

Algorithms of this class recover $\hat{\mathbf{x}}$ by successive projection and thresholding operations. Given some initial approximation $\check{\mathbf{x}}^{(0)}$ to the transform coefficients $\check{\mathbf{x}} = \Psi \mathbf{x}$, the solution is calculated in the following manner:

$$\check{\mathbf{x}}^{(i)} = \check{\mathbf{x}}^{(i)} + \frac{1}{\gamma} \Psi \Phi^T \left(\mathbf{y} - \Phi \Psi^{-1} \check{\mathbf{x}}^{(i)} \right), \qquad (3)$$

$$\check{\mathbf{x}}^{(i+1)} = \begin{cases} \check{\check{\mathbf{x}}}^{(i)}, & \left| \check{\check{\mathbf{x}}}^{(i)} \right| \geq \tau^{(i)}, \\ 0 & \text{else,} \end{cases} \qquad (4)$$

where γ is a scaling factor, and $\tau^{(i)}$ is the threshold used at the ith iteration. Further observation of this process shows us that this procedure is actually a specific instance of a projected Landweber (PL) algorithm [15]. We note that convergence of IHT has been shown in [5].

IHT recovery improves reconstruction speed by at least an order of magnitude and maintains a high degree of accuracy. Reconstruction time can be further reduced by implementing a block-based measurement and recovery procedure, as proposed in [3]. In this technique, Φ is applied on a block-by-block basis, while the reconstruction step incorporates a smoothing operation (such as Weiner filtering) into the IHT. By employing blocking, the results in [3] show a reduction of computation time by four orders of magnitude for comparable accuracy versus linear programming approaches. In [4], this method is referred to as block CS and smoothed PL (BCS-SPL) and is extended via the use of directional transforms. The algorithm in [4] is given as

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

Here, $\mathbf{x}^{(0)} = \mathbf{\Phi}^T \mathbf{y}$. The method uses hard thresholding with a fixed convergence factor λ for all iterations [6], and can be calculated as a function of the number of coefficients used in $\mathbf{\Psi}$ [16].

3. DC-BCS-SPL

In [4], BCS-SPL was shown to be both more computationally efficient and to provide more accurate reconstructions than other recovery techniques, especially when using directional transforms as the sparse basis. We now propose a method which incorporates disparity estimation and compensation as side information into this competitive recovery algorithm with the goal of improving recovery accuracy when used within the multiview setting. We exploit the strong correlations between multiview images by reconstructing the residual between images and their disparity-compensated predictions as a means for refining the accuracy of direct BCS-SPL reconstruction. Our method requires no additional information from the encoder, simply the typical CS formulation—namely, the measurement matrix, $\mathbf{\Phi}_d$; a set of measurements, $\mathbf{y} = \mathbf{\Phi}_d \mathbf{x}_d$; and the sparsity basis, $\mathbf{\Psi}$. We refer to this proposed method as disparity-compensated BCS-SPL (DC-BCS-SPL).

The DC-BCS-SPL algorithm, depicted in Fig. 1, is partitioned into two phases. In the first phase, a prediction of the current image, $\mathbf{x}_d$, is created by bidirectionally interpolating the BCS-SPL reconstructions of the two nearest views (the left and right neighbors), i.e. $\mathbf{x}_p = \text{ImageInterpolation}(\hat{\mathbf{x}}_{d-1}, \hat{\mathbf{x}}_{d+1})$. Next, the residual, $\mathbf{r}$ is calculated by taking the difference between the given measurements, $\mathbf{y}_d$, and the projection of $\mathbf{x}_p$ onto the measurement basis, $\mathbf{y}_p = \mathbf{\Phi}_d \mathbf{x}_p$. This residual, $\mathbf{r} = \mathbf{y}_p - \mathbf{y}$, is then reconstructed using BCS-SPL and added back to $\mathbf{x}_p$ to obtain the reconstruction $\hat{\mathbf{x}}_d$.

In the second phase, the reconstruction obtained from the first phase is used to refine the prediction, $\mathbf{x}_p$. Disparity estimation is used to find two sets of disparity vectors, $\mathbf{DV}_{d-1}$ and $\mathbf{DV}_{d+1}$, between $\hat{\mathbf{x}}_d$ and the reconstructions of its neighbor images. The dis-

parity vectors are then used to produce two disparity-compensated predictions of $\hat{\mathbf{x}}_d$ which are averaged together to produce a single prediction. This prediction will serve as the $\mathbf{x}_p$ for the next reconstruction. This process is repeated k times.

The iterative process improves the quality of the final reconstruction because the use of DC allows us to make a better prediction of the image, which leads to smoother and more easily reconstructed residuals, which then allow us to make more accurate predictions, and so on. DC-BCS-SPL converges quickly—typically iterating for $2 \leq k \leq 5$ is sufficient.

4. EXPERIMENTAL RESULTS

In order to observe the effectiveness of the DC-BCS-SPL recovery, we evaluate the performance of the proposed method against that of the direct-recovery approach, i.e., BCS-SPL used to reconstruct the frame independently of its neighbors. We use several transforms, specifically a DCT, DWT, complex dual-tree DWT (DDWT), and contourlet transform (CT). In our results, we refer to the implementations of the direct approach simply by the name of the used transform, and DC-*transform* is used to refer to the implementations of DC-BCS-SPL using the named transform. In our simulations, disparity vectors are calculated using a full block-based search with integer-pixel accuracy, a block size of 16×16, and a search window of 32×32. It is conceivable that the performance of DC-BCS-SPL could be increased with more sophisticated disparity-vector estimation. For DC-BCS-SPL, we consider two measurement block sizes, 32×32 and 64×64, and the wavelet based transforms are computed to 5 and 6 levels of decomposition, respectively, for these block sizes. Additionally, all images within the measured multiview set are projected using the same subrate.

Tables 1 and 2 present the performance, in PSNR, for several 512×512 images from the Middlebury multiview database[1] at several subrates, M/N, and for the two measurement block sizes considered. All images are rectified, and any radial distortion is removed. It should be noted that, due to the variation in quality that can result from differences in random measurement matrices, all PSNR values represent an average of 5 independent trials.

As illustrated in Fig. 2, the quality of DC-BCS-SPL is overall ∼2 dBs higher than the PSNR performance obtained by using direct BCS-SPL under the same conditions. We have found this performance gain to be true regardless of the sparsity basis, $\mathbf{\Psi}$, used. Note that results in Fig. 2 are calculated by using a single iteration ($k = 1$) of reconstruction. Increasing the number of iterations shows further performance gains.

The DC-BCS-SPL method shows a performance improvement of ∼1 dB to ∼3 dB for lower to higher subrates in comparison to direct BCS-SPL. Of the transforms used, the DDWT gave the best performance for both direct and DC BCS-SPL. Additionally, for images with high variation or texture (such as the "Monopoly" multiview image set), the performance gain of the DC method over direct BCS-SPL is even more pronounced, peaking at ∼4.5 dB. It should also be noted that low-variation images benefited from larger measurement block sizes, as can be seen for the "Plastic" multiview image set which shows a performance gain of ∼1.5 dBs when 64×64 blocks are used instead of 32×32 blocks.

5. CONCLUSIONS

In this paper, we proposed a new method for the CS recovery of multiview images which takes advantage of the high degree of inter-frame correlation which is characteristic of the multiview application. We included side information in the form of disparity
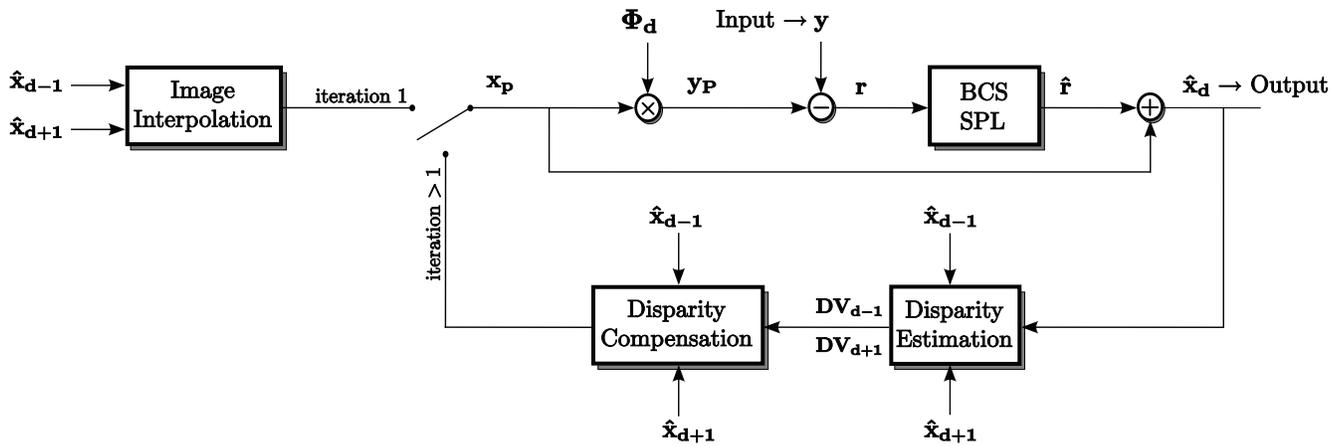

Figure 1: The DC-BCS-SPL reconstruction algorithm.
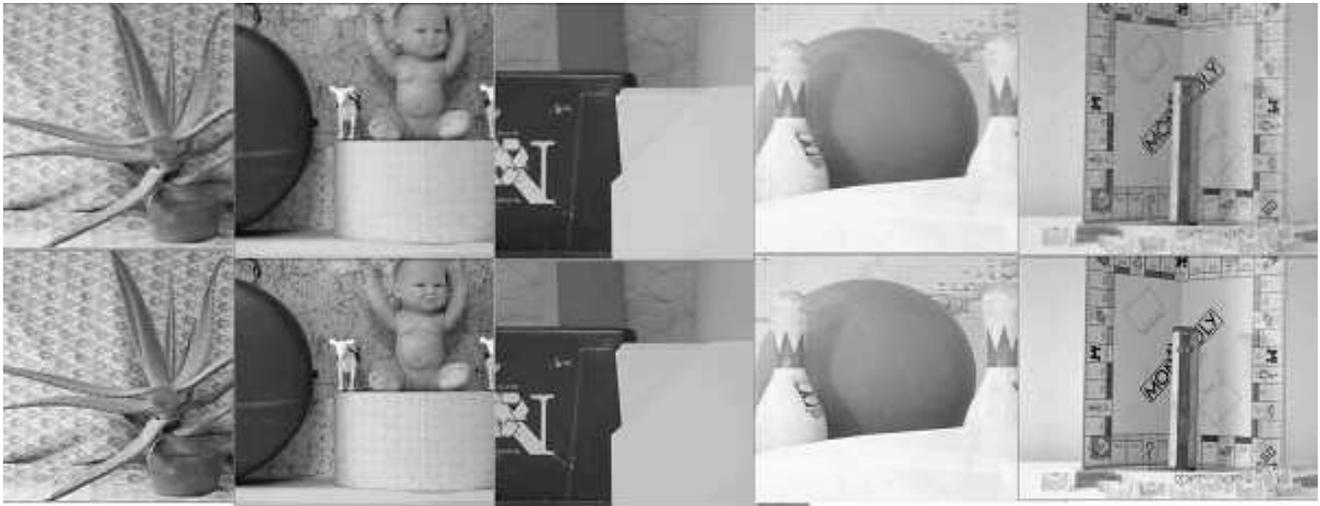
Figure 2: Images from the five multiview sets (left to right: Aloe, Baby, Plastic, Bowling, and Monopoly) reconstructed using the given experimental framework: the first row using direct BCS-SPL, the second row using DC-BCS-SPL

estimation and compensation and using the technique of reconstructing a residual rather than an image, and we incorporated this information into the CS-recovery framework. Experimental results displayed an increase in performance when using this extra information in comparison to recoveries which merely reconstruct each image independently from one another.

6. REFERENCES

[1] 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.

[2] 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, pp. 83–91, March 2008.

[3] 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.

[4] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *Proceedings of the International Conference on Image Processing*, Cairo, Egypt, November 2009, pp. 3021–3024.

[5] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *The Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 629–654, December 2008.

[6] K. K. Herrity, A. C. Gilbert, and J. A. Tropp, "Sparse approximation via iterative thresholding," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Toulouse, France, May 2006, pp. 14–19.

[7] 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*

Table 1: PSNR performance for Aloe, Baby, Plastic, Bowling, Monopoly (512×512, Middlebury database): 32×32 blocksize for BCS-SPL.

Aloe					
Subrate/PSNR (dB)	0.1	0.2	0.3	0.4	0.5
DCT	25.24	26.95	28.23	29.44	30.69
DC-DCT	25.67	28.16	30.00	31.77	33.59
DWT	25.70	27.44	28.91	30.31	31.67
DC-DWT	26.34	29.08	31.16	33.04	34.89
DDWT	25.88	27.68	29.17	30.61	32.07
DC-DDWT	**26.61**	**29.34**	**31.50**	**33.47**	**35.43**
CT	25.88	27.75	29.19	30.56	31.93
DC-CT	26.55	29.27	31.27	33.10	34.91
Baby					
Subrate/PSNR (dB)	0.1	0.2	0.3	0.4	0.5
DCT	30.51	33.16	35.11	36.86	38.60
DC-DCT	31.34	34.65	37.00	39.15	41.30
DWT	30.77	33.61	35.64	37.45	39.25
DC-DWT	31.49	35.53	38.07	40.32	42.52
DDWT	31.00	33.78	35.79	37.60	39.37
DC-DDWT	**32.13**	**35.77**	**38.26**	**40.56**	**42.72**
CT	30.84	33.62	35.63	37.42	39.16
DC-CT	32.12	35.48	37.78	39.86	41.89
Plastic					
Subrate/PSNR (dB)	0.1	0.2	0.3	0.4	0.5
DCT	31.98	35.94	39.12	41.76	44.03
DC-DCT	**32.68**	36.69	40.39	44.26	47.32
DWT	31.58	36.04	39.58	42.64	45.31
DC-DWT	31.57	35.16	38.66	44.32	47.79
DDWT	31.72	36.28	39.88	43.02	45.84
DC-DDWT	31.38	35.24	39.13	44.04	**48.97**
CT	32.03	36.35	39.39	42.05	44.48
DC-CT	31.99	**37.04**	**41.51**	**44.64**	47.07
Bowling					
Subrate/PSNR (dB)	0.1	0.2	0.3	0.4	0.5
DCT	32.41	35.44	37.65	39.79	41.76
DC-DCT	33.33	37.00	39.80	42.10	44.54
DWT	32.60	35.96	38.42	40.61	42.64
DC-DWT	33.36	37.61	40.96	43.46	45.85
DDWT	32.70	36.08	38.61	40.87	42.94
DC-DDWT	33.66	**38.10**	**41.54**	**44.07**	**46.56**
CT	32.55	35.76	38.06	40.20	42.15
DC-CT	**33.74**	37.48	40.31	42.54	44.65
Monopoly					
Subrate/PSNR (dB)	0.1	0.2	0.3	0.4	0.5
DCT	26.34	28.74	31.55	33.78	36.00
DC-DCT	27.95	32.03	34.86	37.82	40.35
DWT	26.15	29.26	31.89	34.34	36.76
DC-DWT	27.29	32.39	36.05	39.20	41.98
DDWT	26.23	29.49	32.28	34.79	37.19
DC-DDWT	27.48	32.82	**36.58**	**39.55**	**42.18**
CT	26.73	29.58	32.10	34.42	36.62
DC-CT	**28.73**	**33.06**	35.99	38.58	40.96

Table 2: PSNR performance for Aloe, Baby, Plastic, Bowling, Monopoly (512×512, Middlebury database): 64×64 blocksize for BCS-SPL.

Aloe						
Subrate/PSNR (dB)	0.05	0.1	0.2	0.3	0.4	0.5
DCT	23.63	25.31	26.96	28.24	29.43	30.70
DC-DCT	24.01	25.81	28.13	29.94	31.66	33.45
DWT	24.45	25.71	27.43	28.88	30.26	31.66
DC-DWT	24.78	26.48	29.10	31.15	33.02	34.90
DDWT	24.58	25.90	27.65	29.14	30.56	32.05
DC-DDWT	**24.89**	**26.66**	**29.33**	**31.48**	**33.45**	**35.44**
CT	24.40	25.90	27.70	29.15	30.51	31.90
DC-CT	24.73	26.61	29.25	31.28	33.10	34.94
Baby						
Subrate/PSNR (dB)	0.05	0.1	0.2	0.3	0.4	0.5
DCT	27.40	30.37	33.05	34.96	36.65	37.63
DC-DCT	28.59	31.34	34.49	36.77	38.71	40.67
DWT	28.97	31.22	33.71	35.67	37.45	39.23
DC-DWT	29.34	32.08	35.61	38.08	40.28	42.47
DDWT	29.13	31.36	33.86	35.81	37.58	39.35
DC-DDWT	29.78	32.31	**35.67**	**38.25**	**40.50**	**42.69**
CT	28.77	31.05	33.63	35.58	37.32	39.03
DC-CT	**29.84**	**32.37**	35.57	37.91	39.99	42.04
Plastic						
Subrate/PSNR (dB)	0.05	0.1	0.2	0.3	0.4	0.5
DCT	30.17	32.72	36.72	38.93	41.09	44.08
DC-DCT	**30.73**	**33.92**	38.89	42.22	44.43	47.73
DWT	28.96	32.66	37.27	40.97	44.10	46.76
DC-DWT	29.70	32.84	39.22	45.10	48.62	50.95
DDWT	29.39	32.84	37.54	41.28	44.47	47.14
DC-DDWT	29.91	33.54	**40.23**	**45.96**	**49.02**	**51.27**
CT	29.87	33.07	37.12	40.19	42.81	45.18
DC-CT	30.01	33.52	39.28	43.35	46.40	49.25
Bowling						
Subrate/PSNR (dB)	0.05	0.1	0.2	0.3	0.4	0.5
DCT	30.55	32.82	35.88	38.23	40.36	41.74
DC-DCT	31.32	34.11	37.65	40.34	42.81	45.09
DWT	30.49	33.36	36.83	39.28	41.38	43.28
DC-DWT	31.26	34.85	39.29	42.24	44.66	46.77
DDWT	30.59	33.45	36.98	39.47	41.58	43.48
DC-DDWT	31.58	**35.21**	**39.71**	**42.60**	**44.94**	**47.06**
CT	30.47	33.11	36.33	38.71	40.78	42.66
DC-CT	**31.59**	34.67	38.59	41.26	43.49	45.72
Monopoly						
Subrate/PSNR (dB)	0.05	0.1	0.2	0.3	0.4	0.5
DCT	24.57	26.55	29.36	31.48	34.26	36.33
DC-DCT	25.42	28.02	32.03	35.13	37.89	39.89
DWT	24.56	26.81	30.19	32.95	35.48	37.85
DC-DWT	24.97	28.25	33.56	**37.13**	**40.00**	**42.64**
DDWT	25.03	27.08	30.23	32.84	35.27	37.55
DC-DDWT	25.24	28.63	**33.58**	36.78	39.50	42.11
CT	24.98	26.93	29.93	32.49	34.79	36.90
DC-CT	**25.97**	**29.09**	33.29	36.30	38.86	41.23

on Selected Areas in Communications, vol. 1, no. 4, pp. 586–597, December 2007.

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

[9] 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.

[10] X. Chen and P. Frossard, "Joint reconstruction of compressed multi-view images," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, April 2009, pp. 1005–1008.

[11] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, November 2009.

[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, pp. 2992–3004, December 2007.

[13] 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.

[14] M. Fornasier and H. Rauhut, "Iterative thresholding algorithms," *Applied and Computational Harmonic Analysis*, vol. 25, no. 2, pp. 187–208, September 2008.

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

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