

DISPARITY-COMPENSATED COMPRESSED-SENSING RECONSTRUCTION FOR MULTIVIEW IMAGES

Maria Trocan[†], Thomas Maugey[‡], James E. Fowler^{*}, Béatrice Pesquet-Popescu[‡]

[†]*Institut Supérieur d'Electronique de Paris*, [‡]*Télécom ParisTech*, ^{*}*Mississippi State University*
maria.trocan@isep.fr, {maugey, beatrice.pesquet}@telecom-paristech.fr, fowler@ece.msstate.edu

ABSTRACT

In a multiview-imaging setting, image-acquisition costs could be substantially diminished if some of the cameras operate at a reduced quality. Compressed sensing is proposed to effectuate such a reduction in image quality wherein certain images are acquired with random measurements at a reduced sampling rate via projection onto a random basis of lower dimension. To recover such projected images, compressed-sensing recovery incorporating disparity compensation is employed. Based on a recent compressed-sensing recovery algorithm for images that couples an iterative projection-based reconstruction with a smoothing step, the proposed algorithm drives image recovery using the projection-domain residual between the random measurements of the image in question and a disparity-based prediction created from adjacent, high-quality images. Experimental results reveal that the disparity-based reconstruction significantly outperforms direct reconstruction using simply the random measurements of the image alone.

Keywords— Compressed sensing, multiview, disparity compensation, directional transforms

1. INTRODUCTION

More and more applications, like 3D reconstruction, creation of virtual environments, surveillance applications, etc., require systems which capture a scene with several cameras. In these cases, the correlation between images is high because they describe the same scene. Compression, restoration, or other data processing should therefore exploit this redundancy in order to improve performance. The correlation between multiview images can be taken into account by estimating the disparity between them, which corresponds to the displacement of an object between the images and which is a quantity related to the object's depth. Since multiview technology is relatively new, the acquisition of the multiview data can be rather costly. However, the acquisition cost of multiview images could be greatly reduced if only some of the multiviews are captured at high resolution or high fidelity; the other views could possibly be acquired at a lower acquisition cost and thereby be reduced in quality. Such lower acquisition cost could be effectuated by using a compressed-sensing

(CS) recovery of these latter images. CS (e.g., [1]) is a recent paradigm which allows describing a signal with a rate lower than Nyquist without any loss. This is possible under a certain hypothesis of sparsity, and is often driven by linear projection onto random basis. Such random-projection-based signal acquisition could feasibly be accomplished using a so-called single-pixel camera [2]; the corresponding reconstruction can be achieved via any one of a number of emerging schemes for CS image reconstruction (e.g., [3, 4, 5]).

In this paper, we propose to incorporate disparity compensation (DC) into the CS reconstruction of multiview images. In [4], an efficient block-based CS reconstruction of images using directional transforms was proposed. Our goal here is to improve the performance of this algorithm by considering disparity information at the reconstruction. The results that we obtain are promising and demonstrate that we can reach a recovery quality of more than 50 dB with an acquisition sampling rate divided by at least two. As previously mentioned, we anticipate that this paradigm can be useful in a multiview acquisition wherein some cameras have lower quality than others.

The remainder of the paper is organized as follows. Sec. 2 gives an overview of the CS paradigm introducing the basics for our method which is in turn presented in detail in Sec. 3. Experimental results demonstrating the efficiency of the DC scheme are presented in Sec. 4. Finally, some concluding remarks are made in Sec. 5.

2. BACKGROUND

In CS, a real-valued signal x of length N has to be recovered from M samples, where $M \ll N$ [1]. In other words, x should be reconstructed from the observations $y = \Phi x$, where y has length M , and $\Phi_{M \times N}$ is called the measurement matrix. This recovery is possible if x is sufficiently sparse in a certain space. The usual choice for the measurement basis Φ is a random matrix; in the following, we assume that Φ is orthonormal such that $\Phi \Phi^T = I$. In general, the sparsity condition for x recovery will exist with respect to some unknown transform Ψ . In this case, the key to CS reconstruction is the production of a sparse set of significant transform coefficients, $\tilde{x} = \Psi x$, and the ideal recovery procedure searches for

the \tilde{x} with the smallest l_0 norm consistent with the observed y . However, as this l_0 optimization is NP-complete, several alternative procedures have been proposed. For example, applying a convex relaxation to the l_0 problem results in an l_1 optimization, as exemplified by basis/matching-pursuit-based algorithms [6, 7, 8]:

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

where Ψ^{-1} represents the inverse transform. Generally, such algorithms could be implemented with linear programming.

Recently, projection-based CS-reconstruction techniques have been proposed [9]. Algorithms of this class recover \tilde{x} by successively projecting and thresholding: the reconstruction starts from some initial approximation $\tilde{x}^{(0)}$, which is further refined in an iterative manner, as in the following:

$$\begin{aligned} \tilde{x}^{(i)} &= \tilde{x}^{(i)} + \frac{\Psi \Phi^T}{\lambda} (y - \Phi \Psi^{-1} \tilde{x}^{(i)}) \\ \tilde{x}^{(i+1)} &= \begin{cases} \tilde{x}^{(i)}, & |\tilde{x}^{(i)}| \geq \tau^{(i)} \\ 0, & \text{otherwise} \end{cases}, \end{aligned} \quad (1)$$

where λ is a scaling factor, and $\tau^{(i)}$ is the threshold used at the i^{th} iteration. It is straightforward to see that this procedure is a specific instance of a projected Landweber (PL) algorithm [10].

In [3], a block-based approach of the above paradigm for the CS recovery of 2D images was proposed. In this technique, the sampling of an image is driven by random matrices applied block-by-block to the image, while the reconstruction is a variant of the PL reconstruction of (1) that incorporates a smoothing operation (e.g. Wiener filtering), ostensibly to eliminate block artifacts due to the block-based sampling. Due to its combination of block-based CS (BCS) sampling and smoothed-PL (SPL) reconstruction, this technique was denoted BCS-SPL in [4]; we adopt this same terminology here. The recovery process in BCS-SPL is iterative—the approximation of the image at iteration $i+1$, $x^{(i+1)}$, is obtained from $x^{(i)}$ as [4]:

$$\begin{aligned} \text{function } x^{(i+1)} &= \text{SPL}(x^{(i)}, y, \Phi_{\text{block}}, \Psi, \lambda) \\ \hat{x}^{(i)} &= \text{Wiener}(x^{(i)}) \\ \text{for each block } j \\ \hat{\hat{x}}_j^{(i)} &= \hat{x}_j^{(i)} + \Phi_{\text{block}}^T (y - \Phi_{\text{block}} \hat{x}_j^{(i)}) \\ \tilde{\tilde{x}}^{(i)} &= \Psi \hat{\hat{x}}^{(i)} \\ \tilde{x}^{(i)} &= \text{Threshold}(\tilde{\tilde{x}}^{(i)}, \lambda) \\ \bar{x}^{(i)} &= \Psi^{-1} \tilde{x}^{(i)} \\ \text{for each block } j \\ x_j^{(i+1)} &= \bar{x}^{(i)} + \Phi_{\text{block}}^T (y - \Phi_{\text{block}} \bar{x}_j^{(i)}) \end{aligned} \quad (2)$$

In [4], the initialization is done as $x^{(0)} = \Phi^T y$, and the reconstruction process is stopped once $|D^{(i+1)} - D^{(i)}| < 10^{-4}$,

where D is defined as the mean squared error (MSE), $D^{(i)} = \frac{1}{\text{block.size}} \|x^{(i)} - \hat{x}^{(i-1)}\|_2^2$, between the i^{th} image reconstruction and the first refinement step at the $(i+1)$ iteration. We note that we employ hard thresholding for the operator $\text{Threshold}(\cdot)$, where the convergence factor λ is fixed for all iterations [11] (specifically, it varies as function of the number of coefficients of Ψ from one transform to another [12]). We note also that the convergence for hard-thresholding algorithms of this nature has been proven in [13].

3. DC-BCS-SPL RECONSTRUCTION

In [4], the BCS-SPL reconstruction originating in [3] was demonstrated to provide effective reconstruction for 2D images when used with directional transforms. In the following, we propose an iterative DC algorithm for the reconstruction of multiview images; this algorithm is based on the BCS-SPL method described in the previous section and incorporates the estimation of and compensation for disparity between the multiple views. Since multiview images are strongly correlated, we exploit this correlation by deploying CS reconstruction on the DC residual. The method assumes the same setup as in [4]; that is, for the current image \mathbf{x}_d , which is the image to be CS-reconstructed, we have the projection/measurement matrix Φ ; the set of observations, $\mathbf{y} = \Phi \mathbf{x}_d$; and the directional transform used in the reconstruction, Ψ . Additionally, to adapt the BCS-SPL algorithm to the multiview scenario, we assume that we know images adjacent to \mathbf{x}_d ; specifically, we know the closest images to the “left” and “right” of \mathbf{x}_d which are \mathbf{x}_{d-1} and \mathbf{x}_{d+1} , respectively.

The DC-BCS-SPL algorithm is partitioned into two phases. In the first phase, a predictor \mathbf{x}_p for \mathbf{x}_d is created by bidirectionally interpolating the closest views, $\mathbf{x}_p = \text{ImageInterpolation}(\mathbf{x}_{d-1}, \mathbf{x}_{d+1})$. Next, we calculate the residual \mathbf{r} between the original observation \mathbf{y} and the observation resulting from the projection of \mathbf{x}_p using the same measurement matrix, Φ . This residual then drives the BCS-SPL reconstruction. We note that, alternatively, \mathbf{x}_p could be given by the direct BCS-SPL reconstruction of the current image, i.e., $\mathbf{x}_p = \text{BCS-SPL}(\mathbf{y}, \Phi, \Psi)$. However, we have found that, at low subrates (M/N small), the quality of the interpolated image is much better than that of the direct BCS-SPL reconstruction.

In the second phase, the reconstructed residual $\hat{\mathbf{r}}$ is further refined with reverse DC to obtain the final reconstruction $\hat{\mathbf{x}}_d$. In the second phase, \mathbf{DV}_{d-1} and \mathbf{DV}_{d+1} are the left and right disparity vectors, respectively; these are obtained from disparity estimation (DE) applied to the current reconstruction, $\hat{\mathbf{x}}_d$, of the current image and the left and right adjacent images. The disparity vectors then drive the DC of the current image to produce the current prediction, \mathbf{x}_p , and its corresponding residual, \mathbf{r} . We note that the second phase of the algorithm is repeated k times. The complete algorithm is

presented below:

Given Φ , Ψ , and $y = \Phi x_d$:

$$(1) \begin{cases} x_p = \text{ImageInterpolation}(x_{d-1}, x_{d+1}) \\ y_p = \Phi x_p \\ r = y - y_p \\ \hat{r} = \text{BCS-SPL}(r, \Phi, \Psi) \\ \hat{x}_d = x_p + \hat{r} \end{cases}$$

Repeat k times:

$$(2) \begin{cases} \{DV_{d-1}, DV_{d+1}\} = DE(\hat{x}_d, x_{d-1}, x_{d+1}) \\ x_p = DC(\hat{x}_d, DV_{d-1}, DV_{d+1}) \\ y_p = \Phi x_p \\ r = y - y_p \\ \hat{r} = \text{BCS-SPL}(r, \Phi, \Psi) \\ \hat{x}_d = x_p + \hat{r} \end{cases}$$

As illustrated in Fig. 1, the quality of DC-based reconstruction is several dBs higher than that obtained by direct BCS-SPL reconstruction. We have found this to be true regardless of the transform Ψ employed. Note that Fig. 1 is for a single iteration ($k = 1$) of phase 2 of the reconstruction; further improvement results from iteratively repeating phase 2. Given the quality of the reconstruction after phase 1, the predictor at each step will be obtained by DC between the current reconstructed image and its neighbors; the improvement in reconstruction quality is due to the refinement of the disparity vectors, leading to a smoother residual at each step which is much easily reconstructed by BCS-SPL.

Note that the original images (x_{d-1} and x_{d+1}) are used as references for DE. This is pertinent, since the proposed algorithm serves to reduce the acquisition cost (camera quality) by at least 25% (equivalent to a subrate $M/N = 0.5$, the maximum we consider). We note also that phase 2 of the proposed algorithm converges quickly—typically, $2 \leq k \leq 5$ is sufficient for convergence in PSNR to the second decimal place.

4. EXPERIMENTAL RESULTS

In this section, we present more comprehensive experimental results, evaluating several directional transforms for both direct and DC-based CS reconstruction. Specifically, we deploy a discrete cosine (DCT), a discrete wavelet (DWT), a dual-tree discrete wavelet (DDWT) [14], and a contourlet transform (CT) [15] within the BCS-SPL framework as described in Sec. 3. We refer to the resulting implementations as *transform* for direct CS reconstruction using the transform in question, and *DC-transform* for the corresponding DC scheme using the algorithm of Sec. 3; here, *transform* $\in \{\text{DCT}, \text{DWT}, \text{DDWT}, \text{CT}\}$. In our simulations, the disparity



(a)



(b)

Fig. 1. Monopoly, 512 \times 512: BCS-SPL reconstruction using 64 \times 64 DCT at subrate $M/N = 0.2$. (a) Direct BCS-SPL (PSNR = 29.03 dB); (b) one-step DC-BCS-SPL (PSNR = 42.70 dB).

is estimated using a full-search block-based DE algorithm, where the size of the block is 16 \times 16, and the search area is 32 \times 32 pixels. For BCS-SPL, we have used a 64 \times 64 block size for the sampling and reconstruction processes. The number of decomposition levels for the tested transforms is 6. We use the BCS-SPL implementation available from its authors¹.

Figs. 2–5 present the PSNR performance for several 512 \times 512 images from the Middlebury database² at several subrates, M/N . All images are rectified and the radial distortion has been removed. It should be noted that, since the quality of reconstruction can vary due to the randomness of the

¹<http://www.ece.msstate.edu/~fowler/>

²<http://cat.middlebury.edu/stereo/data.html>

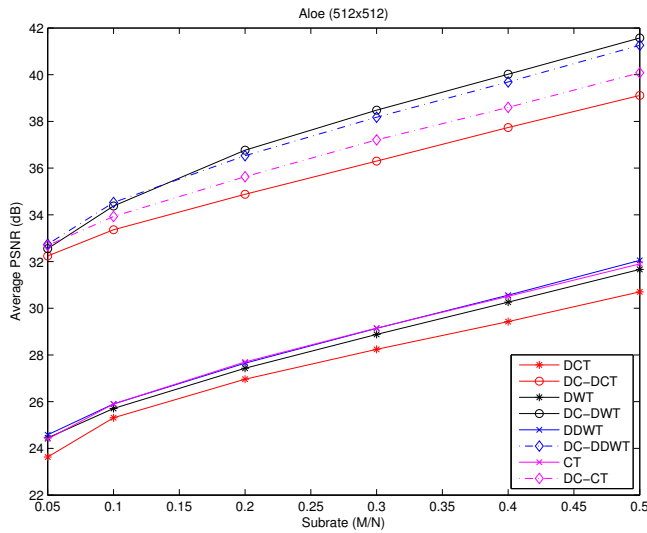


Fig. 2. Reconstruction quality (dB) for "Aloe" test image, as a function of the subrate, and for different transforms.

measurement matrix Φ , all PSNR values in the figures are obtained by averaging 5 independent trials. It is evident that the DC-based recovery leads to higher-quality results, having an average gain of ~ 7 dB with respect to direct BCS-SPL reconstruction. The results confirm that both direct BCS-SPL as well as DC-BCS-SPL with the DDWT achieve the best performance at both low and high subrates. Moreover, for highly textured content (e.g., the Monopoly image), the gain of the DC-based reconstruction over the direct reconstruction reaches a peak of ~ 13 dB.

5. CONCLUSION

In this paper, we have considered the situation in which random projections coupled with CS reconstruction are used to reduce image-acquisition cost within a multiview setting. Specifically, we have assumed that an image is subject to random projections during its acquisition, and that high-quality adjacent images are available to aid its CS reconstruction. We have proposed the incorporation of DE and DC into the CS reconstruction, such that two adjacent images are used to form a prediction of the current image in between them. This predicted image is then projected using the same measurement matrix as was used to acquire the random CS projections of the current image. CS reconstruction then proceeds on the residual between the projected prediction and the projected image. Experimental results reveal a substantial increase in reconstruction quality for the DC-based algorithm as opposed to a simple, direct CS reconstruction driven by the random measurements of the image rather than the projection-domain residual.

We note that, although we have specifically considered the

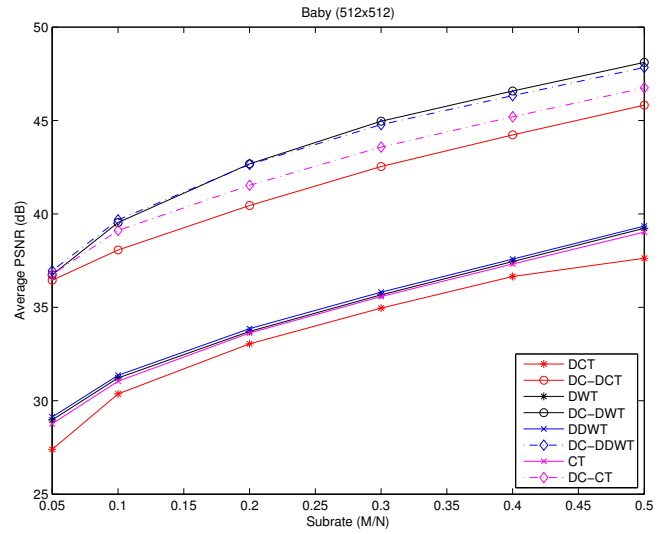


Fig. 3. Reconstruction quality (dB) for "Baby" test image, as a function of the subrate, and for different transforms.

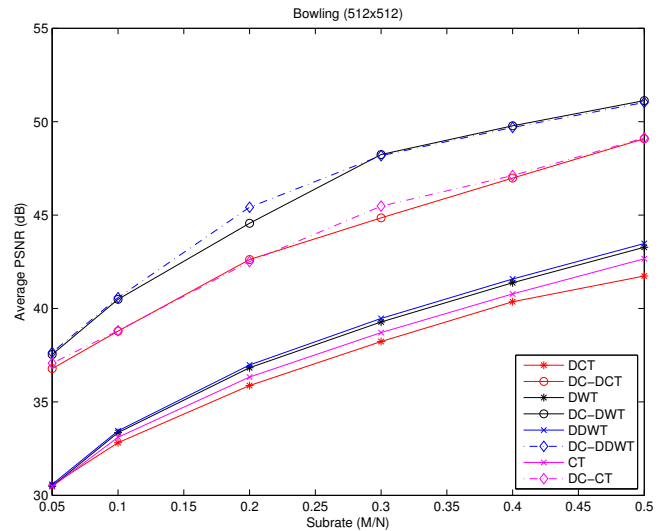


Fig. 4. Reconstruction quality (dB) for "Bowling" test image, as a function of the subrate, and for different transforms.

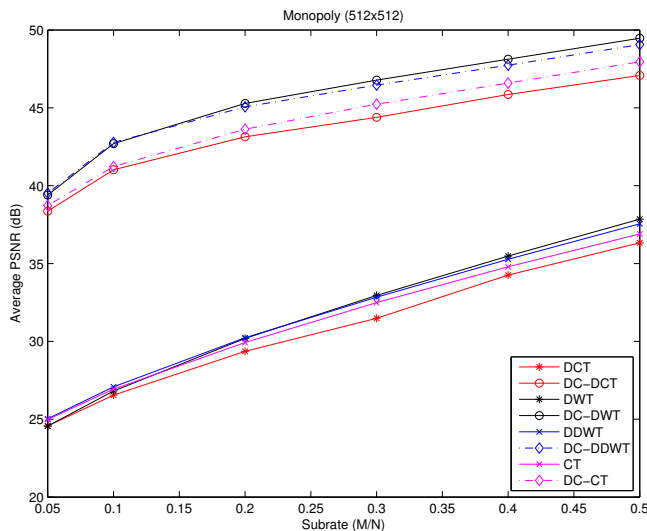


Fig. 5. Reconstruction quality (dB) for "Monopoly" test image, as a function of the substrate, and for different transforms.

multiview setting, we anticipate that the techniques presented here are also applicable to stereo images in which one image is acquired with high quality and the other is subject to CS-based random projections. In the DC-BCS-SPL algorithm we present here, one would simply modify the prediction process so as to be unidirectional rather than bidirectional.

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] 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," in *Computational Imaging IV*, C. A. Bouman, E. L. Miller, and I. Pollak, Eds. San Jose, CA: Proc. SPIE 6065, January 2006, p. 606509.
- [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] 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.
- [6] 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.
- [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 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 42nd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, October 2008, pp. 581–587.
- [9] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Transactions on Information Theory*, vol. 52, no. 49, pp. 4036–4048, 2006.
- [10] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Bristol, UK: Institute of Physics Publishing, 1998.
- [11] 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.
- [12] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [13] 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.
- [14] 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.
- [15] 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.