

BLOCK-WISE LENSLESS COMPRESSIVE CAMERA

Xin Yuan, Gang Huang, Hong Jiang and Paul A. Wilford

Nokia Bell Labs, 600 Montain Avenue, Murray Hill, NJ, 07974, USA

ABSTRACT

The existing lensless compressive camera (L^2C^2) [1] suffers from low capture rates, resulting in low resolution images when acquired over a short time. In this work, we propose a new regime to mitigate these drawbacks. We replace the global-based compressive sensing used in the existing L^2C^2 by the local block (patch) based compressive sensing. We use a single sensor for each block, rather than for the entire image, thus forming a multiple but spatially parallel sensor L^2C^2 . This new camera retains the advantages of existing L^2C^2 while leading to the additional benefits of fast acquisition, real time reconstruction, and the flexibility to image size. We develop multiple geometries of this block-wise L^2C^2 in this paper. We have built prototypes of the proposed block-wise L^2C^2 and demonstrated excellent results of real data.

Index Terms— Compressive sensing, Lensless compressive imaging, denoising, sparse representation, real time.

1. MOTIVATION

Inspired by compressive sensing (CS) [2, 3], diverse compressive cameras [4–17] have been built. The single-pixel camera [4], which implemented the compressive imaging in space, is an elegant architecture to prove the concept of CS. However, the hardware elements used in the single-pixel camera are more expensive than conventional CCD or CMOS cameras, thus limit the applications of this new imaging regime. The lensless compressive camera (L^2C^2), proposed in [1], enjoys low-cost property and has demonstrated excellent results using advanced reconstruction algorithms [18–20].

Though enjoying various advantages, the existing L^2C^2 [1] suffers from low capture rates. Specifically, the currently used sensor can only capture the scene at around 10Hz. For a 64×64 image, if we desire a high resolution image, we need on the order of 10% measurements (relative to the pixel numbers). This requires about 1 minute with current aperture switching technique. This is far from our real time requirement as cameras are built to get instant images/videos. In order to obtain images in a shorter time, we have to sacrifice the spatial resolution, *i.e.* providing a low-resolution image. An alternative solution is to increase the refresh rate of the aperture assembly (Fig. 1). However, even we can achieve a

higher refresh rate using an expensive hardware, the sensor needs a relative long time to integrate the light or a more expensive sensor is required.

There are strong advantages with the L^2C^2 , directly implementing compressive sensing. These include minimizing problems of lenses sensor. But with current approaches, speed is an issue. In this paper, we propose the block-wise L^2C^2 to mitigate this problem.

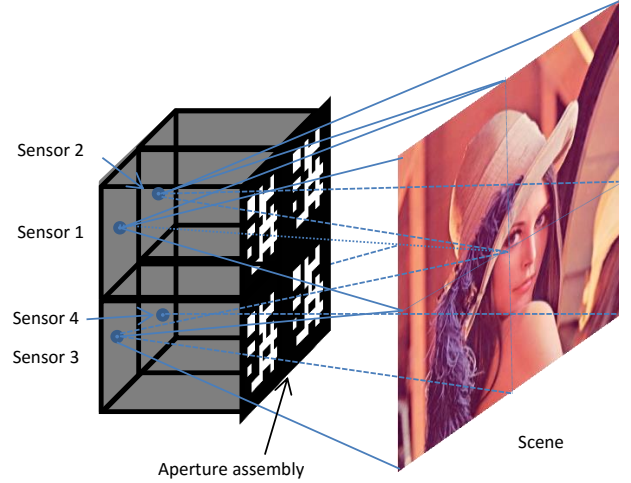


Fig. 1. Geometry and imaging process of the block-wise lensless compressive camera. Four sensors are shown in this example. Each sensor will capture a fraction of the scene. These fractions can be overlapping. The image is reconstructed via first performing patch-based inversion (reconstruction) and then stitching these patches.

2. BLOCK-WISE LENSLESS COMPRESSIVE CAMERA

Figure 1 depicts the geometry of the proposed block-wise L^2C^2 . It consists of three components as shown in Fig. 2: a) the sensor board which contains multiple sensors and each one corresponding to one block, b) the isolation chamber which prevents the light from other blocks, and c) the aperture assembly which can be the same as that used in the L^2C^2 [1].

Note that the block sizes can be different such that each block can reconstruct different resolution image fractions, thus leading to multi-scale compressive camera [21]. The pattern used for each block can also be different and can be adapted to the content of the image part, thus leading to adaptive compressive sensing [22]. In this work, we consider each block uses the same pattern as this will not only save the memory to store these patterns but also enables fast reconstruction for each block [14, 18].

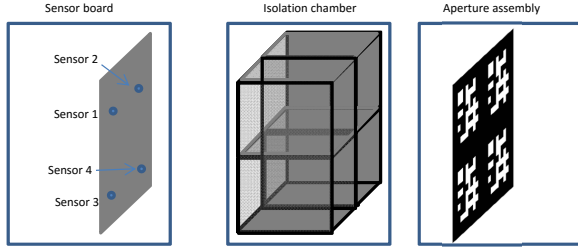


Fig. 2. Components of the block-wise lensless camera. Each parts can be obtained with off-the-shelf components.

This new block-wise L^2C^2 enjoys the following advantages compared to the existing L^2C^2 .

- i) Since each block can be very small, *e.g.* 8×8 pixels, we only need to capture a small number of measurements to achieve high resolution reconstruction. Therefore the capture time can be short.
- ii) The coding patterns used in each block can be the same, therefore the sensing matrix is only of the block size. This saves the memory requirement of sensing matrix as well as speeds up the reconstruction.
- iii) Patch based image reconstruction is fast and since real time stitching algorithms exist, we can perform real time reconstruction.
- iv) These blocks can be integrated to any desirable number, leading to extra high resolution images [21] while retaining the capture rate and fast reconstruction;
- v) The sensor layer can be very close to the aperture assembly, which leads to small size camera. In particular, the thickness of the camera will be extremely small [23].

2.1. Overlapping Regions and Stitching

From the simple geometry shown in Fig. 1, we can see that if the scene is far from the sensor, we will have significant overlapping regions. This will be discussed in next section. On the other hand, the image stitching algorithms are usually based on the features within the overlapping areas to perform registration. Since the angular resolution is limited by the distance between sensors and the aperture assembly, we can adapt this distance in different applications.

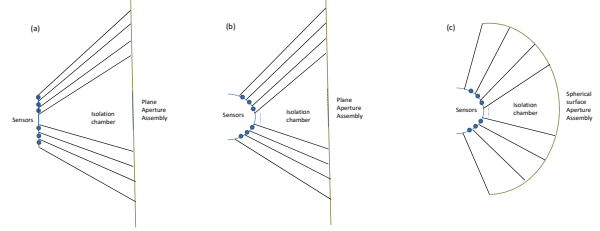


Fig. 3. Cross-sectional view of the “Concentration-Sensor Regime”, where the aperture assembly can be a plane (a-b) or a spherical surface (c). The sensors can be mounted on a plane (a) or a sphere (b-c).

2.2. Concentration-Sensor Regime

In order to mitigate the problem mentioned in Section 2.1, *i.e.* the low angular resolution for far scenes, we propose the following **Concentration-Sensor Regime**, where the sensors are put together in a “cellular” shape. The aperture assembly can be a plane or a spherical surface as shown in Fig. 3. The planer aperture assembly in Fig. 3(a-b) can be replaced by a spherical aperture as shown in Fig. 3(c). The sensor layout of the spherical regime is detailed in Fig. 4. In this configuration, the isolation chamber will be a “trumpet” shape starting from the sensor to the aperture assembly, as demonstrated in Fig. 5. Note the configuration in Fig. 3 (c) and Fig. 4 is a wide angle camera.

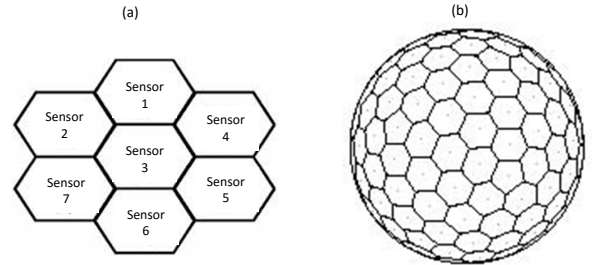


Fig. 4. Sensor layout of the “Concentration-Sensor Regime”. Each sensor covers a hexagon area (a) and the sensor array forms a sphere (b).

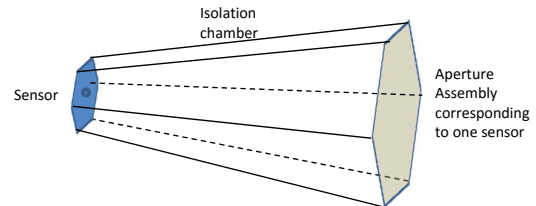


Fig. 5. Isolation chamber in the “Concentration-Sensor Regime”.

3. FAST RECONSTRUCTION

We consider that the patterns used for each block are the same. The measurements can be expressed as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N}, \quad (1)$$

where $\mathbf{X} \in \mathbb{R}^{P \times N_p}$ with P denoting the dimension of the vectorized block (with $\sqrt{P} \times \sqrt{P}$ pixels) and N_p is the number of blocks used in the camera. $\mathbf{Y} \in \mathbb{R}^{M \times P}$ with $M \ll P$ denotes the measurements and each column corresponds to one block (measurements captured by one sensor), and \mathbf{N} signifies the measurement noise. $\mathbf{A} \in \mathbb{R}^{M \times N_p}$ is the sensing matrix, which can be random [24] or the Hadamard matrix with random permutation [19].

3.1. Dictionary Based Inversion

Introducing a basis or (block-based) dictionary \mathbf{D} , equation (1) can be reformulated as

$$\mathbf{Y} = \mathbf{A}\mathbf{D}\mathbf{S} + \mathbf{N}, \quad (2)$$

where $\mathbf{D} \in \mathbb{R}^{P \times Q}$ can be an orthonormal basis with $Q = P$ or an over-complete dictionary [25] with $P \ll Q$. This dictionary can be pre-learned for fast inversion or learned *in situ* [26, 27]. $\mathbf{S} \in \mathbb{R}^{Q \times N_p}$ is desired to be sparse so that various ℓ_1 -based algorithms can be used to solve the following problem [28, 29]

$$\min \|\mathbf{S}\|_1, \quad \text{subject to} \quad \mathbf{Y} = \mathbf{A}\mathbf{D}\mathbf{S}, \quad (3)$$

given \mathbf{A} , \mathbf{Y} and \mathbf{D} or variant problems [30].

Diverse algorithms have been proposed to solve the above problem and we will use the advanced Gaussian mixture model described below since it does not need any iteration as closed-form analytic solution exists.

3.2. Closed-form Inversion via Gaussian Mixture Models

The Gaussian mixture model (GMM) has recently been re-recognized as an efficient dictionary learning algorithm [20, 31–34]. Recall the image blocks $\mathbf{X} \in \mathbb{R}^{P \times N_p}$ extracted from the image. For i -th patch \mathbf{x}_i , it is modeled by a GMM with K Gaussian components [31]:

$$\mathbf{x}_i \sim \sum_{k=1}^K \pi_k \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \quad (4)$$

where $\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$ represent the mean and covariance matrix of k -th Gaussian, and $\{\pi_k\}_{k=1}^K$ denote the weights of these Gaussian components, and $\sum_k \pi_k = 1$.

Dropping the block index i , in a linear model $\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \in \mathcal{N}(\mathbf{0}, \mathbf{R})$, if $\mathbf{x} \sim p(\mathbf{x})$ in (4), then $p(\mathbf{x}|\mathbf{y})$ has the

following analytical form [31]

$$p(\mathbf{x}|\mathbf{y}) = \sum_{k=1}^K \tilde{\pi}_k \mathcal{N}(\mathbf{x}|\tilde{\boldsymbol{\mu}}_k, \tilde{\boldsymbol{\Sigma}}_k) \quad (5)$$

where

$$\tilde{\pi}_k = \frac{\pi_k \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x}_k, \mathbf{R}^{-1} + \mathbf{A}\boldsymbol{\Sigma}_k\mathbf{A}^T)}{\sum_{l=1}^K \pi_l \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x}_l, \mathbf{R}^{-1} + \mathbf{A}\boldsymbol{\Sigma}_l\mathbf{A}^T)} \quad (6)$$

$$\tilde{\boldsymbol{\Sigma}}_k = (\mathbf{A}^T\mathbf{R}\mathbf{A} + \boldsymbol{\Sigma}_k^{-1})^{-1}, \quad (7)$$

$$\tilde{\boldsymbol{\mu}}_k = \tilde{\boldsymbol{\Sigma}}_k(\mathbf{A}^T\mathbf{R}\mathbf{y} + \boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k). \quad (8)$$

While (5) provides a posterior distribution for \mathbf{x} , we obtain the point estimate of $\hat{\mathbf{x}}$ via the posterior mean:

$$\mathbb{E}[\hat{\mathbf{x}}] = \sum_{k=1}^K \tilde{\pi}_k \tilde{\boldsymbol{\mu}}_k. \quad (9)$$

which is a closed-form solution.

Note that $\{\pi_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}_{k=1}^K$ are pre-trained on other datasets and given \mathbf{A} , $\tilde{\boldsymbol{\Sigma}}_k$ only needs to be computed once and saved. Same techniques can be used for $\mathbf{A}\boldsymbol{\Sigma}_k\mathbf{A}^T$. The only computation left for each block is to calculate $\{\tilde{\boldsymbol{\mu}}_k, \tilde{\pi}_k\}$, which can be obtained very efficiently.

Most importantly, no iteration is required using the above GMM method, leading to *real-time* reconstruction for each block. Furthermore, each block can be reconstructed in parallel with GPU. Since real time stitching algorithm exists, after blocks are obtained via the GMM, we can get the entire image instantly. Alternatively, we can pre-train a neural network for the inversion and obtain the reconstruction in real time [35].

4. RESULTS

In this section, we first conduct the simulation to verify the proposed approach and then describe the hardware prototypes we have built to demonstrate real data results.

4.1. Simulation

We consider each block with size 8×8 pixels and reconstruct the images with different numbers of measurements. The image is of size 512×512 . We assume that there is no overlapping among blocks and the image part is ideally captured by different sensors. The PSNR (Peak-Signal-to-Noise-Ratio) of the reconstructed image is used as a metric and plotted in Fig. 6 with exemplar reconstructed images shown in Fig. 7. The GMM is used for reconstruction. We can observe that even a single measurement in each block provides excellent reconstruction. The running time for reconstruction is < 0.01 seconds using an i7 CPU.

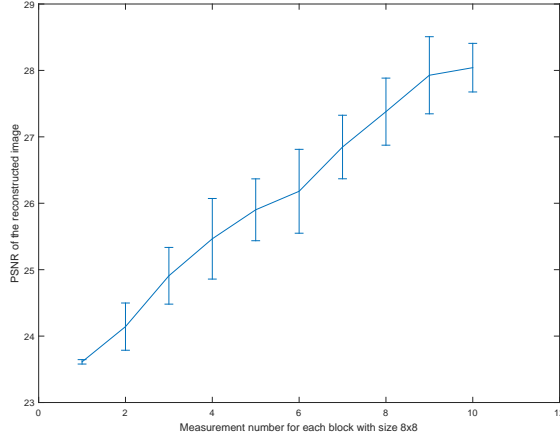


Fig. 6. PSNR of the reconstructed image using different number of measurements for each 8×8 (pixels) block. 10 trials were performed with different random binary coding patterns.



Fig. 7. Example of reconstructed images using different number of measurements (from top-left to bottom-right: 1 to 10) for each 8×8 pixels block.

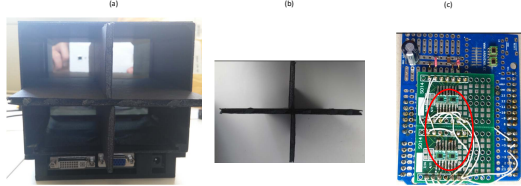


Fig. 8. Prototype with 4 sensors used to prove the concept, (a) LCD used as aperture assembly with the isolation chamber (b) and (c) the 4 sensors.

4.2. Proof of Concept Prototypes and Results

We first used 4 sensors to build the block-wise L^2C^2 to prove the concept, with photos shown in Fig. 8. We used 32×32 pixels pattern for each sensor and these patterns are shared across these 4 sensors. The scene is a photo printed on a paper and pasted on a wall. Without employing any stitching algorithm, one result is shown in Fig. 9, where the CSr is defined as the number of measurements relative to pixels in the reconstructed image. It can be observed that good results can be obtained by only using 10% of the pixel data (CSr = 0.1), which is 102 measurements captured by each sensor.

Next we show the reconstruction results with 16 sensors (a 4×4 array) in Fig. 10, where we restricted the light to test

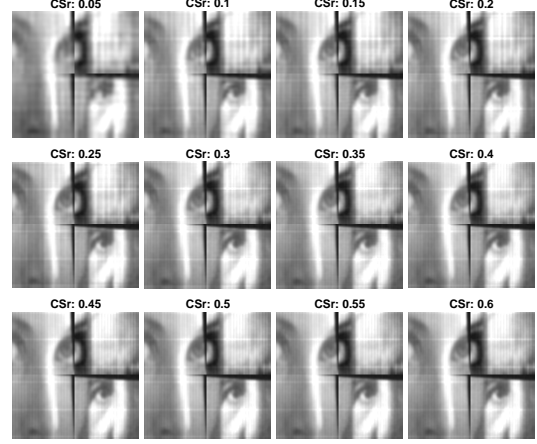


Fig. 9. Results from the measurements taken by the prototype with 4 sensors.

the performance of the camera in dark environment. It can be observed that even when CSr = 0.05 (as each block is of 16×16 pixels, CSr = 0.05 denotes each sensor only captured 12 measurements), we can still get reasonable results.

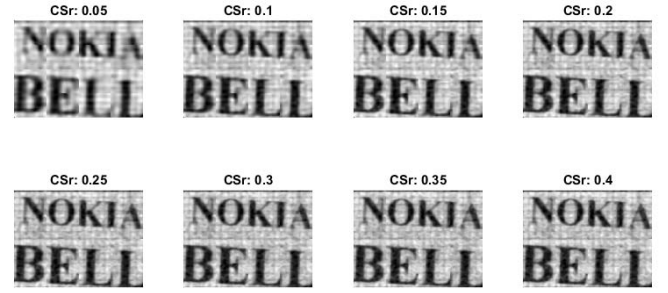


Fig. 10. Results from the measurements taken by the prototype with 16 sensors.

5. CONCLUSIONS

We have proposed the block-wise lensless compressive camera to mitigate the speed issue existing in the current L^2C^2 . Multiple geometries of this block-wise L^2C^2 have been developed. Prototypes have been built to demonstrate the feasibility of the proposed imaging architecture. Excellent results of real data have been shown to verify the fast capture and real time reconstruction of the proposed camera.

6. REFERENCES

- [1] G. Huang, H. Jiang, K. Matthews, and P. Wilford, "Lensless imaging by compressive sensing," *IEEE International Conference on Image Processing*, 2013.
- [2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [3] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

- [4] 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, 2008.
- [5] A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied Optics*, vol. 47, no. 10, pp. B44–B51, 2008.
- [6] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. J. Brady, "Coded aperture compressive temporal imaging," *Optics Express*, pp. 698–706, 2013.
- [7] D. Reddy, A. Veeraraghavan, and R. Chellappa, "P2C2: Programmable pixel compressive camera for high speed imaging," *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [8] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Video from a single coded exposure photograph using a learned over-complete dictionary," in *IEEE International Conference on Computer Vision (ICCV)*, 2011.
- [9] X. Yuan, P. Llull, X. Liao, J. Yang, G. Sapiro, D. J. Brady, and L. Carin, "Low-cost compressive sensing for color video and depth," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [10] T.-H. Tsai, X. Yuan, and D. J. Brady, "Spatial light modulator based color polarization imaging," *Optics Express*, vol. 23, no. 9, pp. 11912–11926, May 2015.
- [11] T.-H. Tsai, P. Llull, X. Yuan, D. J. Brady, and L. Carin, "Spectral-temporal compressive imaging," *Optics Letters*, vol. 40, no. 17, pp. 4054–4057, Sep 2015.
- [12] Y. Sun, X. Yuan, and S. Pang, "High-speed compressive range imaging based on active illumination," *Opt. Express*, vol. 24, no. 20, pp. 22836–22846, Oct 2016.
- [13] X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 95–108, Sept 2016.
- [14] X. Yuan, X. Liao, P. Llull, D. Brady, and L. Carin, "Efficient patch-based approach for compressive depth imaging," *Appl. Opt.*, vol. 55, no. 27, pp. 7556–7564, Sep 2016.
- [15] X. Yuan, Y. Sun, and S. Pang, "Efficient patch-based approach for compressive depth imaging," *Appl. Opt.*, vol. 56, no. 10, pp. 2697–2704, 2017.
- [16] X. Yuan and S. Pang, "Structured illumination temporal compressive microscopy," *Biomedical Optics Express*, vol. 7, pp. 746–758, 2016.
- [17] X. Yuan and S. Pang, "Compressive video microscope via structured illumination," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sept 2016, pp. 1589–1593.
- [18] X. Yuan, H. Jiang, G. Huang, and P. Wilford, "SLOPE: Shrinkage of local overlapping patches estimator for lensless compressive imaging," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 8091–8102, November 2016.
- [19] X. Yuan, H. Jiang, G. Huang, and P. Wilford, "Lensless compressive imaging," *arXiv:1508.03498*, 2015.
- [20] X. Yuan, H. Jiang, G. Huang, and P. Wilford, "Compressive sensing via low-rank Gaussian mixture models," *arXiv:1508.06901*, 2015.
- [21] D. J. Brady, M. E. Gehm, R. A. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, and S. D. Feller, "Multiscale gigapixel photography," *Nature*, no. 486, pp. 386–389, 2012.
- [22] X. Yuan, J. Yang, P. Llull, X. Liao, G. Sapiro, D. J. Brady, and L. Carin, "Adaptive temporal compressive sensing for video," in *2013 IEEE International Conference on Image Processing (ICIP)*, Sept 2013, pp. 14–18.
- [23] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. Baraniuk, "Flatcam: Thin, bare-sensor cameras using coded aperture and computation," *arXiv:1509.00116*, 2015.
- [24] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?," *IEEE Transactions on Information Theory*, 2006.
- [25] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [26] X. Yuan, T.-H. Tsai, R. Zhu, P. Llull, D. J. Brady, and L. Carin, "Compressive hyperspectral imaging with side information," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 964–976, September 2015.
- [27] X. Yuan, "Compressive dynamic range imaging via Bayesian shrinkage dictionary learning," *Optical Engineering*, vol. 55, no. 12, pp. 123110, 2016.
- [28] X. Liao, H. Li, and L. Carin, "Generalized alternating projection for weighted- $\ell_{2,1}$ minimization with applications to model-based compressive sensing," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 797–823, 2014.
- [29] X. Yuan, "Generalized alternating projection based total variation minimization for compressive sensing," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sept 2016, pp. 2539–2543.
- [30] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Img. Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.
- [31] M. Chen, J. Silva, J. Paisley, C. Wang, D. Dunson, and L. Carin, "Compressive sensing on manifolds using a non-parametric mixture of factor analyzers: Algorithm and performance bounds," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6140–6155, December 2010.
- [32] G. Yu, G. Sapiro, and Stéphane Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Transactions on Image Processing*, 2012.
- [33] J. Yang, X. Yuan, X. Liao, P. Llull, G. Sapiro, D. J. Brady, and L. Carin, "Video compressive sensing using Gaussian mixture models," *IEEE Transaction on Image Processing*, vol. 23, no. 11, pp. 4863–4878, November 2014.
- [34] J. Yang, X. Liao, X. Yuan, P. Llull, D. J. Brady, G. Sapiro, and L. Carin, "Compressive sensing by learning a Gaussian mixture model from measurements," *IEEE Transaction on Image Processing*, vol. 24, no. 1, pp. 106–119, January 2015.
- [35] K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "Reconnet: Non-iterative reconstruction of images from compressively sensed random measurements," in *CVPR*, 2016.