2018 5th International Workshop on Compressed Sensing applied to Radar, Multimodal Sensing, and Imaging (CoSeRa)

Single-pixel real-time video imaging with closed-form single-step image reconstruction

Anna Pastuszczak*, Krzysztof M. Czajkowski and Rafał Kotyński
University of Warsaw, Faculty of Physics, Pasteura 5, 02-093 Warsaw, Poland
*Email: apastuszczak@igf.fuw.edu.pl

***Abstract*—We demonstrate real-time single-pixel video imaging from highly undersampled data with the recently proposed Fourier domain regularized inversion method (FDRI). FDRI is a non-iterative image reconstruction method based on the calculation of the generalized inverse of the measurement matrix. The regularization of the solution is obtained by minimizing the sum of norms of the convolutions between the reconstructed image and a set of spatial filters. The FDRI method allows us to reconstruct 256 x 256 images at the frame rate of 11 Hz on the fly during the optical compressive measurement with the compression ratio of 3%. With a precalculated generalized inverse matrix, the numerical cost of image reconstruction is proportional to the number of measured samples and the number of pixels in the measured image. FDRI does not require orthogonality of the sensing matrix and it and is especially efficient for highly incomplete measurements. Moreover, FDRI may be used both when the sampling functions are correlated or uncorrelated with the signal. The latter case, such as randomly generated Bernoulli patterns, is applicable to any unknown sparse signals, as long as the sparsity basis of that signal is known. On the other hand, the correlated sampling protocols use some a-priori predictions of the most significant coefficients of the signal sparse representation to increase the accuracy of incomplete measurements.**

I. Introduction

Single-pixel imaging is an indirect compressive imaging technique, which requires only a single photodetector to capture an image [1]. This idea is especially attractive for imaging at mid and far infrared, as well as at terahertz or millimeter wavelengths, for which high-resolution cameras are either expensive or unavailable [2]–[4]. Methods based on ℓ^1-norm optimization allow to recover the sparse representation of the image from a compressive measurement [5], [6]. This is a way to reduce the overall acquisition time needed for single pixel imaging. Presently, one of the greatest challenges for single-pixel cameras is real-time imaging of a moving scene [7]–[11]. On one hand, the indirect method of measurement by itself limits the video frame rate, as to capture each frame requires switching the state of a spatial light modulator (SLM) as many times as the number of taken image samples. For a compressive measurement this number is a fraction of the pixel resolution of the image. On the other hand, efficient reconstruction algorithms are necessary to reconstruct the video on the fly. The iterative optimization methods, which provide excellent results in compressive sensing and imaging, are usually not fast enough to match the sampling rate of the state-of-the-art SLMs. For this reason, we focus on the closed-form solutions to the inverse problems, which may be cast to the calculation of the Moore-Penrose pseudoinverse. For an underdetermined system of equations the psudoinverse gives the solution with minimal ℓ^2-norm, so it does not tend to find a sparse representation of the image. However, when combined with a sampling protocol, which by itself is an image compression method, the pseudoinverse offers a reasonably good estimate of the measured image [12]–[15]. Several proposals of sampling schemes utilizing some predictions about the sampled scene have been recently studied and applied to single-pixel imaging. They include ordered Walsh-Hadamard patterns [16], sampling with Fourier domain components [10] and patterns designed using deep learning techniques [11]. The major advantage of using psedoinverse for reconstructing images sampled by a single-pixel camera is the low computational cost of this method. With precalculated psedoinverse of the measurement matrix, the reconstruction requires only a single multiplication of a $n \times k$ matrix and a k-length vector, where n is the number of pixels of the measured image and k is the number of sampling functions used in the measurement.

We have recently proposed a novel single-step method of reconstructing a compressively measured image [17]. The method, called Fourier Domain Regularized Inversion (FDRI) method, includes minimization of a quadratic criterion, which tailors the properties of the solution in the Fourier domain. Similarly to the pseudoinverse reconstructions, the FDRI method is expressed in a closed-form and shares the same linear computational cost as a function of the dimension of the measurement. The proposed criterion combines minimization of the image gradient with a penalty-function for high spatial frequencies in the spectral representation of the image.

II. Image reconstruction with Fourier domain regularized inversion method

Let us now recall the principles of the FDRI method. We assume that the measurement of an image performed by a single-pixel camera is described by a system of equations:

$$Mx = y, \tag{1}$$

where x is a vector representing the measured image, consisting of n real-valued pixels, M is a rectangular $k \times n$ measurement matrix, each row of M represents a single sampling function used to probe the image during the measurement, and y is a vector with $k < n$ elements containing the measurements. The image is reconstructed from the measurement y through minimization of a quadratic criterion constrained by Eq. (1),

$$x_0 = \arg\min_x E(x) \text{ subject to } Mx = y. \tag{2}$$

The criterion is defined as a sum on ℓ^2-norms of the convolution of the image with a set of real-valued filters $h^{(p)}$

$$E(x) = \sum_p \alpha^{(p)} \|x * h^{(p)}\|^2, \tag{3}$$

where $\alpha^{(p)} \geq 0$ are weighting coefficients. The solution to the optimization problem from Eq. (2) takes the following closed form:

$$x_0 = Py, \tag{4}$$

with

$$P = BM^*(MBM^*)^{-1}. \tag{5}$$

Let us further denote with carets the matrices and vectors transformed into the Fourier domain and let F stand for the Fourier transform. Then, the Fourier transform of matrix B in Eq. (5), i.e. $\hat{B} = FBF^*$, is a $n \times n$ diagonal matrix, related to the transfer functions of the filters $h^{(p)}$ in the following form:

$$\hat{B}_{i,j} = \frac{\delta_{i,j}}{\sum_p \alpha^{(p)} |\hat{h}_i^{(p)}|^2} \quad \text{and} \quad B = F^* \hat{B} F, \tag{6}$$

where $\hat{h}^{(p)} = Fh^{(p)}F^*$ is the transfer function of the p-th filter. The matrix P in Eq. (4) is a generalised inverse of the measurement matrix M, and it converges to the Moore-Penrose pseudoinverse of matrix M when $\hat{B} = I = B$, i.e. when all Fourier components are transmitted with equal weights by each of filters $h^{(p)}$.

We are using a single-dimensional notation whenever possible, thus images are represented as vectors. However, it should be kept in mind that they are actually two-dimensional (2D) objects. Therefore, F denotes the 2D Fourier transform and all filters $h^{(p)}$ are defined in 2D Fourier domain. This domain is spanned by spatial frequencies $\omega_x \equiv \omega_x(i) \in (-\pi, +\pi)$, and $\omega_y \equiv \omega_y(i) \in (-\pi, +\pi)$. A single vector index i may be assigned to the values of the 2D spatial frequencies in the same way, as it is assigned to the values of pixels in a 2D image. Using this notation, we propose a combination of four filters in the optimization criterion, namely: a penalty filter for high spatial frequencies $\hat{h}^{(1)} = \sqrt{\omega_x^2 + \omega_y^2}$, gradient filters with transfer functions $\hat{h}^{(2)} = \sin(\omega_x)$, $\hat{h}^{(3)} = \sin(\omega_y)$, responsible for minimizing image nonuniformity, and a constant filter $\hat{h}^{(4)} = const$ required to remove singularity from Eq. (6). For this set of filters, matrix $\hat{B}$ takes the form:

$$\hat{B}_{i,j} = \delta_{i,j} \left(\varepsilon + \mu^2 \frac{\omega_x(i)^2 + \omega_y(i)^2}{2\pi^2} + \right.$$
$$\left. + (1-\mu)^2 \Big(\sin^2\big(\omega_x(i)\big) + \sin^2\big(\omega_y(i)\big) \Big) \right)^{-1}, \tag{7}$$

where $\mu \in [0, 1]$ is a coefficient weighting the contribution of specific filters to the solution, and ε is a small constant. In our further results, we use $\varepsilon = 10^{-4}$.

Finally, the FDRI method may be used for real-time single-pixel video acquisition and reconstruction, provided that the reconstruction matrix P is calculated beforehand. Than the reconstruction of each video frame with Eq. (4) takes only nk multiplications. The practical use of this method consists of the following steps:

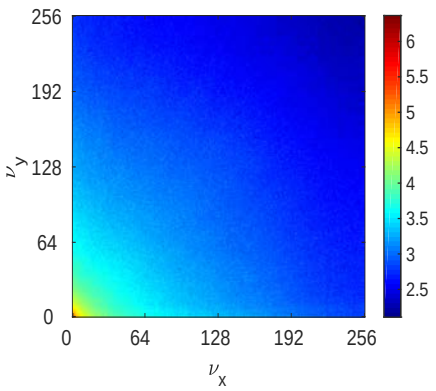

Fig. 1. Averaged 2D DCT spectrum calculated for a database of 85 images with various content at the resolution 256×256 (logarithmic scale).

- Preparation before the measurement:
  1. Select the sampling functions and create the measurement matrix M. In the case of highly compressive measurement (i.e. when $k \ll n$), the best performance is obtained when the Fourier spectrum of the sampling functions is similar to the Fourier spectrum of the measured image.
  2. Select filters $h^{(p)}$ for the optimization criterion from Eq. (3).
  3. Calculate matrix $\hat{B}$ using Eq. (6). With the filters proposed in this study, matrix $\hat{B}$ takes the form given by Eq. (7).
  4. Calculate the reconstruction matrix P using Eq. (5). To reduce the computational cost, matrix B should not be calculated directly. Instead, in Eq. (5) B should be replaced with $F^*\hat{B}F$, and then the direct or inverse fast Fourier transform may be used to calculate the products with F or F^*, respectively.
- Real-time video imaging:
  For each video frame:
  1. Sample the image with the functions stored in the measurement matrix M and measure the elements of vector y.
  2. Reconstruct the image using Eq. (4) and the precalculated matrix P. Both tasks may be conducted in parallel.

III. Numerical and Experimental Results

Let us now demonstrate the performance of the proposed reconstruction method. For this purpose, we consider two qualitatively different types of image sampling, namely the Bernoulli sampling and the discrete cosine transform (DCT) sampling. The Bernoulli sampling consists of randomly generated binary sampling patterns with independent and identically distributed pixel values. This kind of sampling is highly incoherent with most of the standard image sparsity bases, which makes it a preferable choice for classic compressive sensing techniques [6]. The DCT sampling, on the other hand, takes advantage of the prediction of the expected DCT spectrum of the scene. Indeed, for most real-world images,

TABLE I. COMPARISON OF PSNR (AVERAGED OVER 50 TEST IMAGES WITH RESOLUTION 256×256) OBTAINED FOR THREE RECONSTRUCTION METHODS: THE FOURIER DOMAIN REGULARIZED INVERSION METHOD (FDRI), PSEUDOINVERSE-BASED METHOD AND TV MINIMIZATION WITH NESTA. TWO METHODS OF SAMPLING ARE CONSIDERED: RANDOMLY GENERATED BERNOULLI PATTERNS AND BINARIZED ELEMENTS OF THE 2D DCT BASIS SELECTED ACCORDING TO FIG. 2D. IN BOTH CASES, THE COMPRESSION RATIO $k/n \approx 3\%$.

	average PSNR [dB]		
	Pseudoinverse	**FDRI**	**NESTA**
Bernoulli sampling	12.30	16.75	16.21
DCT sampling	20.82	21.81	21.96

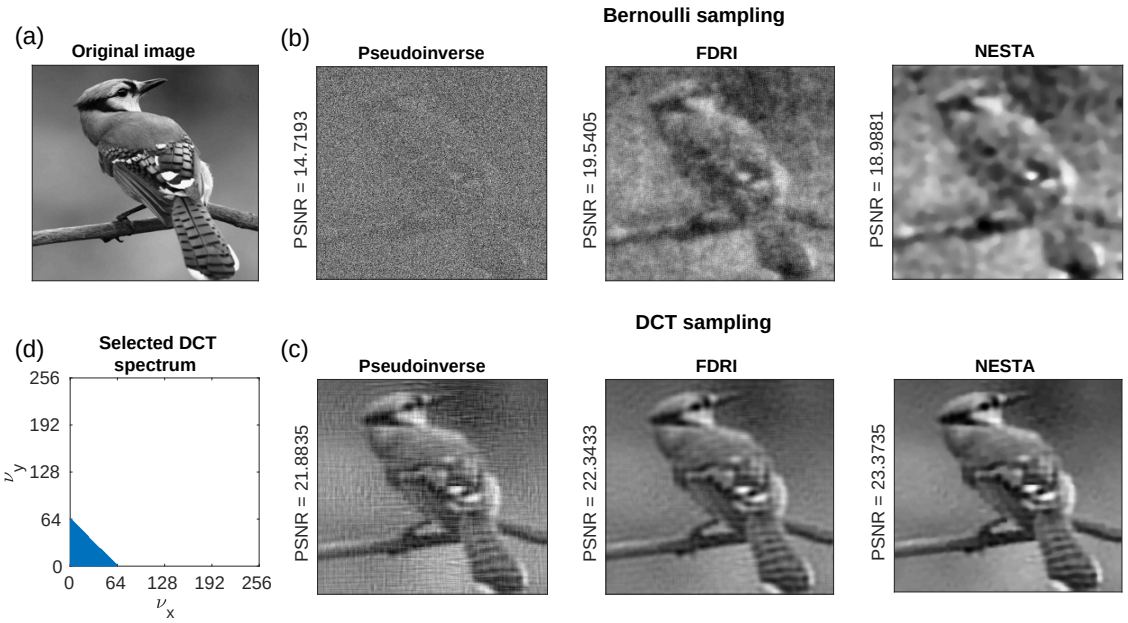

Fig. 2. Comparison of reconstructions obtained for three reconstruction methods: the Fourier domain regularized inversion method (FDRI), pseudoinverse-based method and TV minimization with NESTA. The original image (a) is sampled with 2000 binary patterns with resolution 256×256 (compression ratio $\approx 3\%$). Two types of sampling are considered: (b) incoherent sampling with random patterns generated with Bernoulli distribution and (c) binarized rows from the 2D discrete cosine transform matrix (DCT) corresponding to the low spectral components of the DCT. (d) The range of selected sampling functions from the 2D DCT spectrum (marked in blue).

the 2D DCT spectrum is concentrated around the lowest frequency components. To illustrate this, in Fig. 1 we present the averaged DCT spectrum calculated for a database of 85 images with resolution 256×256, including photographs of faces, landscapes, machines, animals, etc. The DCT sampling is designed to be correlated with the scene. Therefore the sampling matrix is composed of a small set of binarized low-frequency functions selected from the 2D DCT basis. The spectral range of selected DCT functions is depicted in Fig. 2d, however binarization broadens this spectrum to include also some higher frequency components. For both types of sampling the measurement matrix consists of $k = 2000$ binary patterns with resolution 256×256 (the compression ratio $k/n \approx 3\%$).

In Table I we compare the peak signal-to-noise ratio (PSNR) of reconstructions obtained using the FDRI method with reconstructions obtained either with the use of Moore-Penrose pseudoinverse of the measurement matrix or by optimizing the total variation (TV) of the image using NESTA solver. The PSNR values are averaged over a database of 50 images sampled using either Bernoulli or DCT sampling. Exemplary reconstructions obtained with each of the reconstruction methods and sampling types are illustrated in Fig. 2. In most cases, the reconstructions obtained with FRDI and TV minimization methods are of comparable qualities. Both of these methods offer significant improvement over the solutions obtained with the pseudoinverse. While for the DCT sampling, marginally higher PSNR is usually obtained with NESTA, the FDRI method leads to slightly better results when Bernoulli sampling is applied. However, the most practical advantage of the FDRI method over NESTA and other solvers performing the optimization of the image TV or ℓ_1-norm, is the reconstruction time. In the discussed example, the reconstruction of a single image with the FDRI method takes approximately 0.08s on a medium-class PC with single precision arithmetic. This value scales proportionally to the number of sampling functions k and the number of pixels in measured image n. In the same conditions, the TV minimization with NESTA requires several seconds of computations. This makes the FDRI method excellent for applications in single-pixel video imaging with reconstructions performed in real time.

It is also worth mentioning, that the time-resolution of any video signal sampled with a single-pixel camera is proportional to the number of sampling patterns k included in the measure-

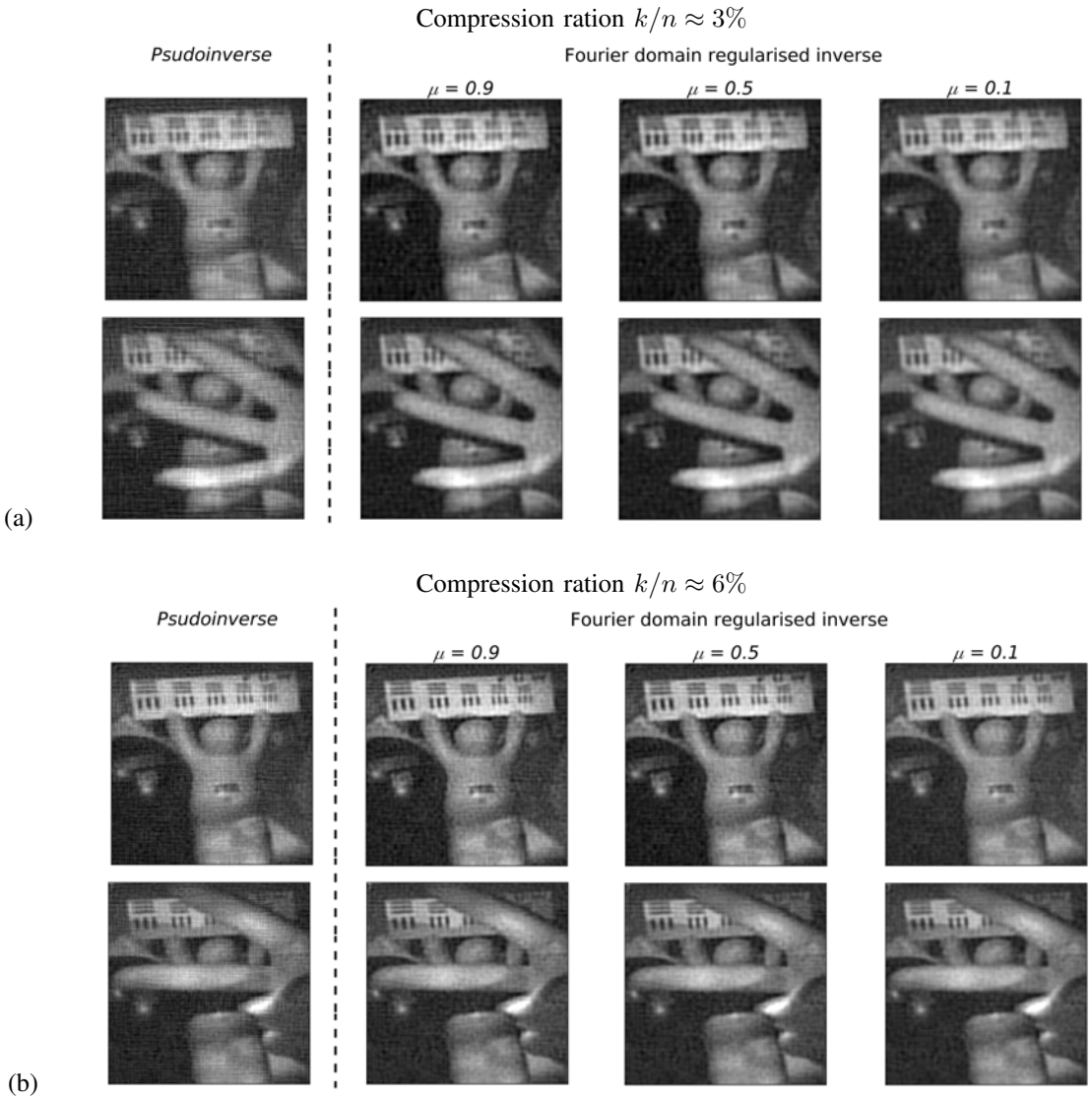

Fig. 3. Experimental reconstructions of selected video frames with resolution 256×256 sampled with a single-pixel camera at compression ratios: (a) $k/n = 3\%$ and (b) $k/n = 6\%$. The images are reconstructed using FDRI method using three different values of coefficient μ correspond to different proportions of specific spectral filters included in the optimization criterion (see Eq. 7). Pseudoinverse reconstructions of the same frames are included for comparison.

ment matrix. Therefore, achieving reasonable frame rate of a single-pixel video requires using efficient sampling methods, which offer high quality of the reconstructions at possibly low compression ratio. The incoherent sampling methods, such as Bernoulli matrices, scrambled Walsh-Hadamard patterns [18] or discrete noiselets [19], allow to uniformly sample a large class of images over their sparsity domain. However, in order to provide high-quality reconstructions, they require the number of sensing patterns k to exceed a certain threshold, which is usually larger than the number of significant coefficients in the image sparse representation [6]. On the other hand, the correlated sampling methods, in which the sampling matrix consists of selected elements from DCT, Fourier or Walsh-Hadamard basis, offer significantly better performance for highly incomplete measurement with $k \ll n$ (see Fig. 2). It should be, however, kept in mind, that correlated sampling is equivalent to lossy image compression, and may be outperformed by incoherent sampling at sufficiently large values of k.

Let us now demonstrate the results obtained with the FDRI method applied to real-time imaging with our experimental single-pixel camera. The camera makes use of a digital micromirror device (DMD) to sample a scene at the rate of 22 kHz. Then the brightness of the modulated scene is measured with a pair of photodiodes using the balanced photodetection technique [17], [20], [21]. Finally, a digital oscilloscope is used to sample and digitize the signals. The results are streamed to a PC for further processing and reconstructing the video images on the fly using the FDRI method. The time required to capture a single video frame with resolution 256×256 at compression ratio $k/n = 3\%$ is approximately 0.09s. Together with the fast single-step method of image reconstruction, the camera allows for real-time video imaging at the frame rate of over 11 Hz. The reconstructions of selected video frames obtained with DCT sampling for two different values of the compression ratio $k/n = 3\%$ and $k/n = 6\%$ are presented in Fig. 3. We also demonstrate the influence of changing the weights between filters $h^{(p)}$

included in the calculations of the reconstruction matrix P. For $\mu = 0.9$ the optimization criterion for the FDRI is dominated by the penalty function for high spatial frequencies, while for $\mu = 0.1$ the gradient filters become dominant (see Eq. (7)). The gradient filters tend to smooth the nonuniformities of the image, while boosting some of the high-frequency Fourier components responsible for preserving the sharpness of edges. However, they are not efficient for clearing any grid-oriented artifacts. The penalty filter for high spatial frequencies is not affected by the orientation of the artifacts, since its spectrum is axially symmetric. However, it usually introduces some level of blurring into the image.

IV. Conclusion

To conclude, we have demonstrated that the FDRI method significantly improves the quality of reconstructions for compressively sampled images, as compared to reconstructions based on the Moore-Penrose pseudoinverse. It allows to obtain values of PSNR comparable with the ones achieved by methods based on TV minimization for both incoherent and correlated sampling methods. On the other hand, FDRI is a closed-form non-iterative method with computational cost linearly dependent on the size of the measurement matrix, i.e. $\mathcal{O}(nk)$. Thanks to this property, the FDRI can be directly applied to real-time reconstructions of video images captured alive with a single-pixel camera. We have demonstrated, that the FDRI method is capable of keeping pace with a DMD, which samples the video image at the rate of 22 kHz using patterns with resolution 256×256.

Acknowledgment

The authors acknowledge partial financial support from the National Science Center, Poland (grant No. UMO-2017/27/B/ST7/00885).

References

[1] M. Duarte, M. Davenport, D. Takbar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-Pixel Imaging via Compressive Sampling," *IEEE Sign. Process. Mag.*, vol. 25, no. 4, pp. 83–91, 2008.

[2] C. M. Watts, C. C. Nadell, J. Montoya, S. Krishna, and W. J. Padilla, "Frequency-division-multiplexed single-pixel imaging with metamaterials," *Optica*, vol. 3, no. 2, pp. 133–138, Feb 2016. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-3-2-133

[3] C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nat. Photonics*, vol. 8, p. 605, jun 2014. [Online]. Available: https://www.nature.com/articles/nphoton.2014.139#supplementary-information

[4] L. Qiao, Y. Wang, Z. Shen, Z. Zhao, and Z. Chen, "Compressive sensing for direct millimeter-wave holographic imaging," *Appl. Opt.*, vol. 54, no. 11, pp. 3280–3289, Apr 2015. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-54-11-3280

[5] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, 2008.

[6] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Probl.*, vol. 23, p. 969, 2007.

[7] M. P. Edgar, G. M. Gibson, R. W. Bowman, B. Sun, N. Radwell, K. J. Mitchell, S. S. Welsh, and M. J. Padgett, "Simultaneous real-time visible and infrared video with single-pixel detectors," *Sci. Rep.*, vol. 5, p. 10669, may 2015. [Online]. Available: https://www.nature.com/articles/srep10669#supplementary-information

[8] G. M. Gibson, B. Sun, M. P. Edgar, D. B. Phillips, N. Hempler, G. T. Maker, G. P. A. Malcolm, and M. J. Padgett, "Real-time imaging of methane gas leaks using a single-pixel camera," *Opt. Express*, vol. 25, no. 4, pp. 2998–3005, Feb 2017. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-25-4-2998

[9] N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox, "Single-pixel optical camera for video rate ultrasonic imaging," *Optica*, vol. 3, no. 1, pp. 26–29, Jan 2016. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-3-1-26

[10] Zhang Zibang, Wang Xueying, Zheng Guoan, and Zhong Jingang, "Fast Fourier single-pixel imaging via binary illumination," *Sci. Rep.*, vol. 7, no. 1, p. 12029, 2017.

[11] C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Sci. Rep.*, vol. 8, no. 1, p. 2369, 2018.

[12] C. Zhang, S. Guo, J. Cao, J. Guan, and F. Gao, "Object reconstitution using pseudo-inverse for ghost imaging," *Opt. Express*, vol. 22, no. 24, pp. 30 063–30 073, 2014.

[13] W. Gong, "High-resolution pseudo-inverse ghost imaging," *Photon. Res.*, vol. 3, p. 234, 2015.

[14] K. M. Czajkowski, A. Pastuszczak, and R. Kotynski, "Single-pixel imaging with Morlet wavelet correlated random patterns," *Sci. Rep.*, vol. 8, p. 466, 2018.

[15] I. Dokmanic, M. Kolundzija, and M. Vetterli, "Beyond Moore-Penrose: Sparse pseudoinverse," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6526–6530.

[16] Sun, M. J., Meng, L. T., Edgar, M. P., Padgett, M. J., and Radwell, N., "A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging," *Sci. Rep.*, vol. 7, no. 1, p. 3464, 2017.

[17] K. M. Czajkowski, A. Pastuszczak, and R. Kotyński, "Real-time single-pixel video imaging with Fourier domain regularization," *Opt. Express*, vol. 26, no. 16, pp. 20 009–20 022, Aug 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-16-20009

[18] L. Gan, T. T. Do, and T. D. Tran, "Fast compressive imaging using scrambled block hadamard ensemble," in *2008 16th European Signal Processing Conference*, Aug 2008, pp. 1–5.

[19] A. Pastuszczak, B. Szczygieł, M. Mikołajczyk, and R. Kotyński, "Efficient adaptation of complex-valued noiselet sensing matrices for compressed single-pixel imaging," *Appl. Opt.*, vol. 55, no. 19, pp. 5141–5148, Jul 2016. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-55-19-5141

[20] F. Soldevila, P. Clemente, E. Tajahuerce, N. Uribe-Patarroyo, P. Andrés, and J. Lancis, "Computational imaging with a balanced detector," *Sci. Rep.*, vol. 6, p. 29181, jun 2016. [Online]. Available: https://www.nature.com/articles/srep29181#supplementary-information

[21] W.-K. Yu, X.-F. Liu, X.-R. Yao, C. Wang, Y. Zhai, and G.-J. Zhai, "Complementary compressive imaging for the telescopic system," *Sci. Rep.*, vol. 4, p. 5834, jul 2014.