Efficient single pixel imaging in Fourier space

2016 J. Opt. 18 085704

(http://iopscience.iop.org/2040-8986/18/8/085704)

IOP Publishing Journal of Optics
J. Opt. 18 (2016) 085704 (6pp) doi:10.1088/2040-8978/18/8/085704

Efficient single pixel imaging in Fourier space

Liheng Bian, Jinli Suo, Xuemei Hu, Feng Chen and Qionghai Dai

Department of Automation, Tsinghua University, Beijing 100084, People's Republic of China

E-mail: jlsuo@tsinghua.edu.cn

Received 24 April 2016, revised 14 June 2016
Accepted for publication 15 June 2016
Published 14 July 2016

Abstract

Single pixel imaging (SPI) is a novel technique capturing 2D images using a bucket detector with a high signal-to-noise ratio, wide spectrum range and low cost. Conventional SPI projects random illumination patterns to randomly and uniformly sample the entire scene's information. Determined by Nyquist sampling theory, SPI needs either numerous projections or high computation cost to reconstruct the target scene, especially for high-resolution cases. To address this issue, we propose an efficient single pixel imaging technique (eSPI), which instead projects sinusoidal patterns for importance sampling of the target scene's spatial spectrum in Fourier space. Specifically, utilizing the centrosymmetric conjugation and sparsity priors of natural images' spatial spectra, eSPI sequentially projects two $\frac{\pi}{2}$-phase-shifted sinusoidal patterns to obtain each Fourier coefficient in the most informative spatial frequency bands. eSPI can reduce requisite patterns by two orders of magnitude compared to conventional SPI, which helps a lot for fast and high-resolution SPI.

Keywords: single pixel imaging, computational ghost imaging, sinusoidal modulation, importance sampling

(Some figures may appear in colour only in the online journal)

1. Introduction

Single pixel imaging (SPI) [1] is a novel incoherent imaging technique. It captures 2D images using a bucket detector instead of array sensors. SPI shares the same imaging scheme with computational ghost imaging [2], which uses a spatial light modulator (SLM) to generate programmable illumination patterns onto the target scene, and uses a bucket detector to collect the correlated lights. Then the 2D scene can be retrieved from the illumination patterns and corresponding 1D single pixel measurements, using either linear correlation methods [3–6] or compressive sensing (CS) techniques [7, 8]. Due to its high signal-to-noise ratio, wide spectrum range, low cost and flexible light-path configuration, SPI has been widely applied in various fields [9–12].

Despite the above advantages over conventional imaging techniques using array sensors, SPI needs numerous illumination patterns to reconstruct an image, which makes it time consuming and memory demanding [13]. Such a large number of patterns is caused by the utilized random modulation, which randomly and uniformly samples all of the target scene's information with no discrimination. Determined by Nyquist sampling theory, it needs at least N measurements to reconstruct an N-pixel image. In particular, more measurements are needed in real applications to compensate for the system noise and the influences from other external factors. As a reference, Sun *et al* [14] used around 10^6 patterns (20 times that of the image pixels) to reconstruct a 256 × 192-pixel image with sufficient quality for subsequent 3D imaging. Although one can utilize compressive sensing [7] to reduce projections, this largely increases computation complexity [8]. Instead of using random patterns, the technique recently proposed in [15] utilizes sinusoidal modulation to sample the scene's information in Fourier space. Specifically, it projects four $\frac{\pi}{2}$-phase-shifted patterns to sample each spatial frequency of the scene's spatial spectrum, and saves a lot of projections compared to conventional SPI.

From the statistics [16], most information of natural images is concentrated in low spatial frequency bands and exhibits strong sparsity in Fourier space, as shown in figure 1(a) where several exemplar images and their spatial spectra are presented. This motivates us to utilize an

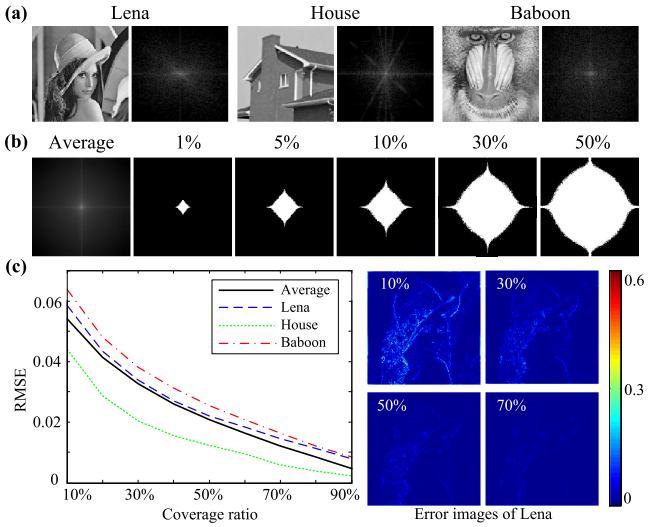

Figure 1. Statistical study of natural images' spatial spectra. (a) Three exemplar natural images and their spatial spectra. (b) The average spatial spectrum of the USC-SIPI database, as well as different acquisition bands under different coverage ratios. (c) The relationship between reconstruction error and coverage ratio.

importance sampling strategy for efficient acquisition in Fourier space, namely that we only capture the most informative spatial frequency bands. To realize the non-uniform sampling of the scene's spatial spectrum, we calculate the statistical importance distribution of nature images' spatial frequencies and sample them in a descending order of importance. To sample each spatial frequency, since random patterns do not work anymore, we use a two-step sinusoidal illumination modulation strategy similar to [17], which is based on the centrosymmetric conjugation property of natural images' spatial spectra. To conclude, we propose an efficient single pixel imaging technique (eSPI) in this paper. The technique utilizes the sparsity and conjugation priors of natural images' spatial spectra to realize fast SPI with extremely high efficiency and low computation cost. We note that the proposed eSPI differentiates from [15] in two aspects: (i) utilizing the sparsity prior of natural images' spatial spectra, eSPI performs importance sampling in the Fourier domain, i.e., the eSPI does not sample all of the Fourier coefficients exhaustively as [15]; (ii) incorporating the centrosymmetric conjugation property of natural images' spatial spectra into the patterning strategy, eSPI needs only two sinusoidal $\frac{\pi}{2}$-phase-shifted patterns for each spatial frequency, instead of four as in [15]. Benefitting from these two strategies, eSPI can save most projections of [15]. In the following, we begin to introduce eSPI in two steps.

2. Methods

The first step of eSPI is to determine the acquisition band in Fourier space, i.e., to decide which Fourier coefficients to sample. Here we first study the statistical distribution of natural scenes' spatial spectra, and accordingly determine the priority of spectrum sampling. Specifically, we transform all 44 images in the USC-SIPI common miscellaneous database [18] to Fourier space, and calculate the spectra's average magnitude map, as shown in the first image in figure 1(b). Then we threshold it to determine the acquisition bands under different coverage ratios (the ratio between the acquisition band and the whole spectrum). The results are shown in figure 1(b), where the white areas stand for the acquisition bands.

Based on the thresholding results, one can determine the acquisition band by setting different coverage ratios according to specific applications. Larger coverage ratio results in a wider acquisition band and more detailed information, but more projections. To further study the relationship between coverage ratio and reconstruction error, we successively sample the spatial spectrum of each image in the above dataset under different coverage ratios, transform them back to spatial space, and calculate reconstruction errors in terms of root-mean-square error (RMSE). RMSE is defined as $\sqrt{E((I_1 - I_2)^2)}$ to measure the difference between two images I_1 and I_2, where E is the pixel-wise average operation. The average performance is plotted as the black solid curve in figure 1(c), where reconstruction errors of several exemplar images are also plotted with dashed lines. The results indicate that though different images are of slight diversity, they follow the same trend that reconstruction error decreases as coverage ratio increases. Besides, the reconstruction residues of the 'Lena' image at different coverage ratios are also presented on the right side as a reference.

After the acquisition band determined, we move on to the second step of eSPI, i.e., sampling each Fourier coefficient in the band to perform non-uniform acquisition. Since random patterns do not work anymore, we use a two-step sinusoidal illumination modulation strategy similar to [17] based on the centrosymmetric conjugation property of natural images' spatial spectra. To introduce the illumination patterning strategy in detail, we first analyze the information encoded by the single pixel measurements in Fourier space. According to the Fourier theorem, a 2D image I can be represented as $I = \sum_i c_i B_i$, where B_i is the ith normalized Fourier basis, and c_i is its Fourier coefficient. Similarly, by applying Fourier transform to a projected pattern P, we can get $P = \sum_j \hat{c}_j B_j$. Its corresponding single pixel measurement s can be represented as

$$\begin{aligned} s &= \left| \sum_m \sum_n I(m, n) P(m, n) \right| \\ &= \left| \sum_m \sum_n \left[\sum_i c_i B_i(m, n) \right] \left[\sum_j \hat{c}_j B_j(m, n) \right] \right| \\ &= \left| \sum_i \sum_j c_i \hat{c}_j \left[\sum_m \sum_n B_i(m, n) B_j(m, n) \right] \right|. \end{aligned} \tag{1}$$

$$s = \Sigma\Sigma(\quad I \quad \odot \quad P \quad)$$

$$= \Sigma\Sigma(\quad \odot \quad)$$

(Spatial domain)

$$= \Sigma\Sigma(\quad \odot \quad)$$

(Fourier domain)

$$= 2(a_0a_1 - b_0b_1) + d_0d_1 \ .$$

Figure 2. Illustration of the encoded information in a correlated single pixel measurement when a real valued sinusoidal pattern is projected. © Playboy Enterprises, Inc.

Here (m, n) index the 2D spatial coordinate. Substituting the orthogonality of Fourier bases

$$f(x) = \begin{cases} \sum_m \sum_n B_i(m, n)B_j(m, n) = 0, & i \neq j \\ \sum_m \sum_n B_i(m, n)B_j(m, n) = 1, & i = j, \end{cases} \tag{2}$$

into the above equation, we get

$$s = \left| \sum_j c_j \hat{c}_j \right|. \tag{3}$$

From this we can see that $\{\hat{c}_j\}$ is a spectrum sampling vector to record the scene's Fourier coefficients. Therefore, we can directly sample a specific Fourier coefficient by setting $\{\hat{c}_j\}$ as a delta vector (containing only one non-zero entry), which results in a sinusoidal pattern with complex values.

However, real facilities can only project real-valued sinusoidal patterns, each having three non-zero coefficients in its spatial spectrum—two conjugate coefficients of a centrosymmetric non-zero frequency pair and one of zero frequency. The conjugation property also holds for natural images. Let $c_1 = a_0 + jb_0$, $c_2 = a_0 - jb_0$ and $c_3 = d_0$ (j is the imaginary unit) denote the three non-zero coefficients of the target scene I, and $\hat{c}_1 = a_1 + jb_1$, $\hat{c}_2 = a_1 - jb_1$ and $\hat{c}_3 = d_1$ represent corresponding coefficients of a sinusoidal pattern P, we have

$$\begin{aligned} s &= |c_1\hat{c}_1 + c_2\hat{c}_2 + c_3\hat{c}_3| \\ &= |(a_0 + jb_0)(a_1 + jb_1) + (a_0 - jb_0)(a_1 - jb_1) + d_0d_1| \\ &= 2(a_0a_1 - b_0b_1) + d_0d_1. \end{aligned} \tag{4}$$

A more explicit demonstration is shown in figure 2. Note that if the pattern's pixel number in each dimension is even, determined by the symmetry property of discrete Fourier transform, there is no corresponding centrosymmetric counterpart of the highest spatial frequency, i.e., the highest frequency cannot form a conjugation frequency pair.

Based on the above derivations, acquiring a specific Fourier coefficient turns into computing a_0 and b_0, with s, a_1, b_1 and d_1 known. To achieve this, we sequentially project three patterns onto the target scene. The first one is a uniform pattern with the constant intensity equal to the mean pixel value of P, and corresponding measurement is exactly d_0d_1. The other two patterns are sinusoidal patterns with the Fourier coefficients being $\{a_1 = \frac{1}{2}, b_1 = 0, d_1 = 1\}$ and $\{a_1 = 0, b_1 = \frac{1}{2}, d_1 = 1\}$, respectively. Thus we can obtain a_0 and b_0 by simply subtracting d_0d_1 from the correlated measurements.

Following the above method, we can sample all the Fourier coefficients of the pre-determined acquisition band, by sequentially projecting corresponding sinusoidal patterns (the uniform pattern needs to be projected only once for all the frequencies). Then, the target scene can be recovered by inverse Fourier transform to the obtained spatial spectrum.

3. Results

To validate the proposed eSPI technique, we first conduct a simulation experiment to compare the reconstruction performance of different SPI methods. We set the 'Lena' image (128×128 and 256×256 pixels respectively) as the latent target scene image, and synthesize the measurements of different patterns following equation (1). We set the coverage ratio being 0.1 and 0.3 (corresponding acquisition bands are shown as the fourth and fifth subfigures in figure 1(b)), respectively. The experiment is conducted using Matlab on an Intel i7 3.6 GHz CPU computer, with 16G RAM and 64 bit Windows 7 system. For comparison, the linear correlation based reconstruction method [4, 5] and the compressive sensing based technique [1] are applied on the same set of sinusoidal patterns, as well as the same number of random patterns. Also, we compare eSPI with conventional SPI in the sense of the same speckle transverse size [12] (same spatial frequency), by truncating conventional random patterns' spatial spectra with the same acquisition band (figure 1(b)) as eSPI. The results are shown in figure 3 and table 1. Note that we omit the results of 'S+Linear', since the eSPI reconstruction (namely inverse Fourier transform) is essentially a linear combination of Fourier bases, which is intrinsically the same as the linear correlation based method in the case of sinusoidal patterns.

From both the visual and quantitative results, we can clearly see that eSPI largely outperforms conventional SPI in terms of both efficiency and reconstruction quality. The advantages come from the utilized sparse information encoding strategy. For conventional SPI, the spatial spectra of random patterns are also random. They sample and multiplex the target scene's whole spatial spectrum randomly and uniformly with no discrimination. Thus conventional SPI can not utilize the importance sampling strategy, and needs much more projections for demulplexing and reconstruction. Instead, each sinusoidal pattern in eSPI only encodes a Fourier coefficient pair of the scene's spatial spectrum. Based

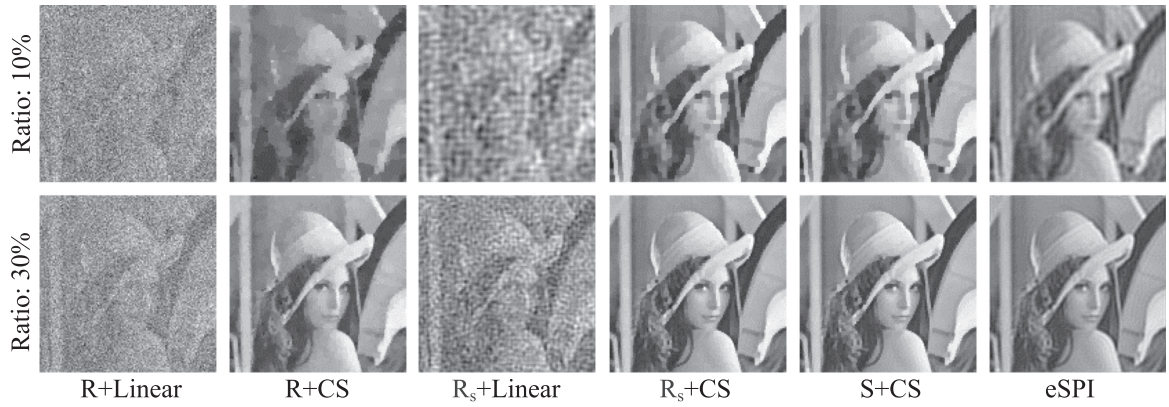

Figure 3. Simulated reconstruction results of the Lena image (128 × 128 pixels) by different SPI strategies, with the coverage ratio being 10% and 30%, respectively. 'R', 'R_s', 'S', 'Linear' and 'CS' stand for random modulation, random modulation of the same speckle traverse size (same spatial frequency) as eSPI, sinusoidal modulation, linear correlation reconstruction, and compressive sensing reconstruction, respectively. © Playboy Enterprises, Inc.

Table 1. Quantitative comparison among different SPI strategies under different coverage ratios and image sizes. The '×' symbol means that the reconstruction is out of memory.

		Ratio: 10%		Ratio: 30%	
		RMSE	Time	RMSE	Time
128 × 128 pixels	R+Linear	0.215	2 s	0.191	6 s
	R+CS	0.115	68 min	0.042	92 min
	R_s+Linear	0.203	2 s	0.187	6 s
	R_s+CS	0.075	68 min	0.041	91 min
	S+CS	0.066	67 min	**0.037**	92 min
	eSPI	**0.061**	**1s**	0.044	**3s**
256 × 256 pixels	R+Linear	0.211	9 s	0.188	26 s
	R+CS	×	×	×	×
	R_s+Linear	0.205	9 s	0.186	25 s
	R_s+CS	×	×	×	×
	S+CS	×	×	×	×
	eSPI	**0.035**	**3s**	**0.014**	**8s**

on this, eSPI samples only the most informative bands and omits unimportant ones. Therefore, it is much more efficient. Note that though the compressive sensing (CS) method produces similar results as eSPI when using sinusoidal patterns, it is much more time consuming and memory demanding. In particular, when image size grows large enough, CS does not work anymore. This is because CS models the reconstruction as an ill-posed problem, which needs large memory and long time for computation under an optimization framework. Instead, eSPI is linear correlation based and does not involve any complex calculations, so it is much faster and saves memory.

To further validate eSPI, we build a proof-of-concept prototype exhibited in figure 4(a). The system mainly consists of two parts including programmable illumination and detection. The illumination part includes a commercial projector's illumination module (numerical aperture of the projector lens is 0.27) and a digital micromirror device (DMD, Texas Instrument DLP Discovery 4100 Development Kit, .7XGA) for spatial modulation. We use the 8-bit mode of the DMD to generate patterns, with the frame rate being 30 Hz. Patterns of 128 × 128 pixels are sequentially projected onto a printed transmissive film (34 mm × 34 mm) as the target scene. Then the correlated lights are recorded by a high-speed bucket detector (Thorlabs DET100 Silicon photodiode, 340–1100 nm) with a 14-bit acquisition board ART PCI8514. The sampling rate is set as 10 kHz. We utilize the self-synchronization technique in [19] to synchronize the DMD and the detector. For each pattern, we average all its corresponding stable measurements for subsequent reconstruction. The coverage ratio of the acquisition band is set as 10%, resulting in 1635 illumination patterns in total. The reconstructed results of two different scenes are shown in figure 4(b), from which we can see that 10% of the pixel number patterns yield satisfying results. Compared to [14] where the requisite pattern number is 20 times that of the pixel number, eSPI can reduce projections by two orders of magnitude. Note that there exist some artifacts in the reconstructed images. This may be caused by several factors, including film glare, light flicker (voltage fluctuation), ambient light, modulation deviation of the DMD, thermal noise of the detector, and so on. Further efforts are needed to address

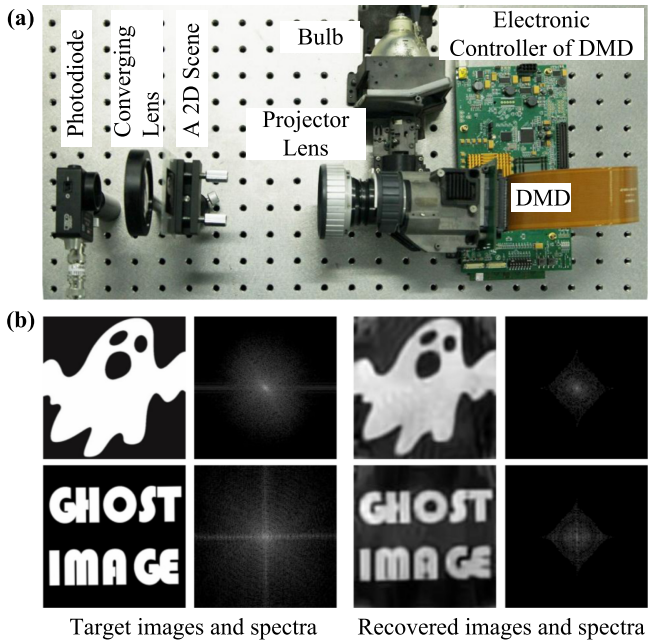

Figure 4. Experiment on real captured data. (a) The eSPI prototype. (b) Reconstruction results of two different scenes (each having 128 × 128 pixels) with the coverage ratio being 10%. The left two columns are the ground-truth target images and their spatial spectra, and the right two columns are corresponding reconstruction.

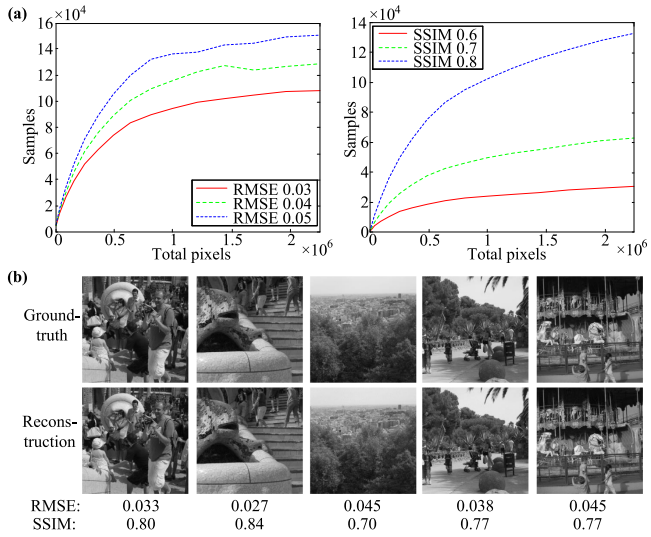

Figure 5. Demonstration of eSPI's advantages for high resolution imaging. (a) Required samplings at different image sizes for the same reconstruction quality. (d) Exemplar reconstructed megapixel images using 10^5 samplings.

these problems by improving the experimental environment and imaging elements, and proposing noise-robust reconstruction techniques.

4. Conclusion and discussion

In this paper, we propose an efficient single pixel imaging technique (eSPI). Different from conventional random illumination modulation which randomly and uniformly samples the scene's whole spatial spectrum, eSPI uses a two-step sinusoidal illumination modulation strategy to obtain the Fourier coefficients of the target scene's most informative spatial frequency bands. As a result, we can reduce the requisite patterns by two orders of magnitude. This helps a lot for fast and high resolution SPI.

Due to the utilized importance sampling strategy, eSPI owns more advantages when applied to high resolution imaging, where the images' spatial spectra are more sparse. To demonstrate this, we downsample each of the 322 natural images (2268 × 1512 pixels) in the Barcelona Calibrated Images Database [20] to different image sizes, and successively sample their spatial spectra under different coverage ratios. Then we transform them back to spatial space, and quantify the reconstruction quality in terms of RMSE and the structure similarity index (SSIM) [21]. SSIM measures the structural similarity between two images. It ranges from 0 to 1, with a larger amount meaning a more similar structure. As shown in figure 5(a), the required sampling number increases slower for the same reconstruction quality as the image size grows. This means that for high-resolution imaging, linearly increased samplings are unnecessary. Specifically, around 10^5 samplings are enough to retrieve a megapixel image with satisfying visual quality, as shown in figure 5(b). We want to note that the low sampling frequencies of eSPI are not caused by the hardware limit. Instead, it is determined by the utilized importance sampling strategy for much higher efficiency with no degeneration of final reconstruction.

eSPI can be widely extended. Since the measurement formation in equation (1) is linear, we can adopt multiplexing [22] to raise the signal-to-noise ratio of final reconstruction. Besides, the content-adaptive sampling scheme [23] can be introduced for higher efficiency. In addition, there exist many other generative image representation methods such as discrete cosine transform. It is interesting to study the pros and cons by applying these transforms to the proposed eSPI framework. What is more, as the requisite number of illumination patterns is largely reduced, eSPI offers promising potentials for real time SPI. These are our future work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 61327902 and 61120106003).

References

[1] Duarte M F, Davenport M A, Takhar D, Laska J N, Sun T, Kelly K E, Baraniuk R G *et al* 2008 Single-pixel imaging via compressive sampling *IEEE Signal Proc. Mag* **25** 83
[2] Shapiro J H 2008 Computational ghost imaging *Phys. Rev.* A **78** 061802
[3] Bromberg Y, Katz O and Silberberg Y 2009 Ghost imaging with a single detector *Phys. Rev.* A **79** 053840
[4] Gong W and Han S 2010 A method to improve the visibility of ghost images obtained by thermal light *Phys. Lett.* A **374** 1005–8

[5] Ferri F, Magatti D, Lugiato L A and Gatti A 2010 Differential ghost imaging *Phys. Rev. Lett.* **104** 253603
[6] Sun B, Welsh S S, Edgar M P, Shapiro J H and Padgett M J 2012 Normalized ghost imaging *Opt. Express* **20** 16892–901
[7] Katz O, Bromberg Y and Silberberg Y 2009 Compressive ghost imaging *Appl. Phys. Lett.* **95** 131110
[8] Aßmann M and Bayer M 2013 Compressive adaptive computational ghost imaging *Sci. Rep.* **3** 1545
[9] Zhao C, Gong W, Chen M, Li E, Wang H, Xu W and Han S 2012 Ghost imaging lidar via sparsity constraints *Appl. Phys. Lett.* **101** 141123
[10] Hu X, Suo J, Yue T, Bian L and Dai Q 2015 Patch-primitive driven compressive ghost imaging *Opt. Express* **23** 11092–104
[11] Bian L, Suo J, Situ G, Li Z, Fan J, Chen F and Dai Q 2016 Multispectral imaging using a single bucket detector *Sci. Rep.* **6** 24752
[12] Gong W, Zhao C, Yu H, Chen M, Xu W and Han S 2016 Three-dimensional ghost imaging lidar via sparsity constraint *Sci. Rep.* **6** 26133
[13] Erkmen B I and Shapiro J H 2009 Signal-to-noise ratio of gaussian-state ghost imaging *Phys. Rev.* A **79** 023833
[14] Sun B, Edgar M P, Bowman R, Vittert L E, Welsh S, Bowman A and Padgett M J 2013 3D computational imaging with single-pixel detectors *Science* **340** 844–7
[15] Zhang Z, Ma X and Zhong J 2015 Single-pixel imaging by means of Fourier spectrum acquisition *Nat. Commun.* **6** 6225
[16] Marcellin M W 2002 *JPEG2000: Image Compression Fundamentals, Standards and Practice* vol 1 (Springer)
[17] Khamoushi S M M, Nosrati Y and Tavassoli S H 2015 Sinusoidal ghost imaging *Opt. Lett.* **40** 3452–5
[18] Univeisity of Southern California SIPI Image Database http://sipi.usc.edu/database/ [online; accessed 1-Feb-2015]
[19] Suo J, Bian L, Xiao Y, Wang Y, Zhang L and Dai Q 2015 A self-synchronized high speed computational ghost imaging system: a leap towards dynamic capturing *Opt. Laser Technol.* **74** 65–71
[20] Computer Vision Center in Universitat Autònoma de Barcelona. The barcelona calibrated images database. http://cvc.uab.es/color_calibration/Database.html [online; accessed 10-May-2015]
[21] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE T. Image Process.* **13** 600–12
[22] Schechner Y Y, Nayar S K and Belhumeur P N 2007 Multiplexing for optimal lighting *IEEE T. Pattern Anal.* **29** 1339–54
[23] Bian L, Suo J, Situ G, Zheng G, Chen F and Dai Q 2014 Content adaptive illumination for Fourier ptychography *Opt. Lett.* **39** 6648–51