Journal of Modern Optics

ISSN: 0950-0340 (Print) 1362-3044 (Online) Journal homepage: https://www.tandfonline.com/loi/tmop20

Coloured computational imaging with single-pixel detectors based on a 2D discrete cosine transform

Bao-Lei Liu, Zhao-Hua Yang, Xia Liu & Ling-An Wu

To cite this article: Bao-Lei Liu, Zhao-Hua Yang, Xia Liu & Ling-An Wu (2017) Coloured computational imaging with single-pixel detectors based on a 2D discrete cosine transform, Journal of Modern Optics, 64:3, 259-264, DOI: 10.1080/09500340.2016.1229507

To link to this article: https://doi.org/10.1080/09500340.2016.1229507

JOURNAL OF MODERN OPTICS, 2017
VOL. 64, NO. 3, 259–264
http://dx.doi.org/10.1080/09500340.2016.1229507

Taylor & Francis
Taylor & Francis Group

Coloured computational imaging with single-pixel detectors based on a 2D discrete cosine transform

Bao-Lei Liu[a], Zhao-Hua Yang[a], Xia Liu[a] and Ling-An Wu[b]

[a]School of Instrument Science and Optoelectronics Engineering, Beihang University, Beijing, China; [b]Institute of Physics and Beijing National Laboratory for Condensed Matter Physics, Chinese Academy of Sciences, Beijing, China

ABSTRACT

We propose and demonstrate a computational imaging technique that uses structured illumination based on a two-dimensional discrete cosine transform to perform imaging with a single-pixel detector. A scene is illuminated by a projector with two sets of orthogonal patterns, then by applying an inverse cosine transform to the spectra obtained from the single-pixel detector a full-colour image is retrieved. This technique can retrieve an image from sub-Nyquist measurements, and the background noise is easily cancelled to give excellent image quality. Moreover, the experimental set-up is very simple.

ARTICLE HISTORY
Received 4 May 2016
Accepted 19 August 2016

KEYWORDS
Computational imaging; ghost imaging; image detection systems

1. Introduction

Ghost imaging (GI), a non-local imaging method whereby an object is reconstructed by means of intensity correlation between the object beam and a corresponding reference beam has attracted considerable attention in recent years. The first such experiment was performed using entangled photon pairs obtained by spontaneous parametric downconversion (*1*). Later, GI was realized using pseudothermal light (*2–7*) and true thermal light (*8–10*). Though the light sources may be different, conventional GI requires a beam splitter to produce correlated light fields in two separate beams, one beam passing via the object to a bucket detector collecting only the total light intensity, and the other measured directly by a detector with spatial resolution.

Computational GI, proposed in 2008 (*11*), eliminated the need for a beam splitter by a spatial light modulator capable of creating deterministic speckle patterns to illuminate the object (*12,13*). This system is more practical and suitable for remote sensing since the high resolution detector in the reference beam is replaced by a computer-generated propagating field. In GI a ghost image of the object is retrieved by correlating the object beam with the reference beam, but in computational GI the measurements recorded by the bucket (single-pixel) detector in the object beam are convoluted with precomputed intensity distribution patterns.

When combined with compressive sensing (also known as compressive sampling) (*14–16*), GI can reconstruct an image from the data sampled at sub-Nyquist frequencies, and can retrieve an image consisting of N2 pixels using much fewer than N2 measurements with a single-pixel detector (*17*). Compressed sensing exploits the redundancy in the structure of most natural objects to reduce the number of measurements required, and is widely applied in the field of computational imaging.

Recently, a new computational imaging technique was proposed by Zhang et al. (*18*), which can reconstruct high-quality images by acquiring their Fourier spectrum. Employing an approach called four-step phase-shifting sinusoidal illumination, and after illuminating a scene with four two-dimensional (2D) sinusoid patterns, each coefficient of the scene's Fourier spectrum is acquired by four responses. The image is then reconstructed by applying an inverse Fourier transform. This is a compressive sampling-like approach, since most natural images are sparse in the Fourier domain. For image reconstruction, only the light intensities collected by the bucket detector need to be considered, without the need to correlate with the precomputed illumination patterns, which are recovered in the inverse Fourier processing.

In this paper, we propose a different full-colour computational imaging method using a single-pixel detector based on a 2D discrete cosine transform (DCT), which

CONTACT Zhao-Hua Yang yangzh@buaa.edu.cn

is simpler and requires fewer measurements. The spectral coefficients can be obtained after illuminating the scene with only two groups of orthogonal sinusoidal patterns, allowing each coefficient to be acquired from just two responses. The image can then be retrieved by applying an inverse cosine transform to the spectrum acquired. Since this transform provides the same features of energy compaction as a discrete Fourier transform for most natural images (*19*), our technique is also a compressive sampling-like approach, which means an image can be reconstructed from only a few coefficients of the spectrum collected by the single-pixel detector.

2. Theory

The object or target scene is illuminated by a projector with two sets of orthogonal sinusoidal patterns, and the reflected field intensity is collected by a single-pixel detector. Every coefficient of the image's cosine transform spectrum is acquired from the two detector responses corresponding to the two patterns. The scene is retrieved by applying the inverse DCT algorithm to the cosine transform spectrum. The final image's signal-to-noise ratio is very good, since any unwanted background illumination is automatically cancelled, as will be shown below.

The cosine transform employed by Ahmed et al. (*20*) expresses a finite sequence of data points in terms of a sum of cosine functions of different frequencies, and has found wide application in transform image coding. It is the foundation of the JPEG standard for still image coding and the MPEG standard for moving images.

The 2D DCT and inverse DCT of an image array are defined in series form as (*19*)

$$F(u,v) = \frac{2}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} C(u)C(v)f(x,y) \cos\frac{(2x+1)u\pi}{2M} \cos\frac{(2y+1)v\pi}{2N} \tag{1}$$

$$I(x,y) = \frac{2}{\sqrt{MN}} \sum_{u=1}^{M-1} \sum_{v=1}^{N-1} C(u)C(v)F(u,v) \cos\frac{(2x+1)u\pi}{2M} \cos\frac{(2y+1)v\pi}{2N} \tag{2}$$

Here, $C(0) = (2)^{-1/2}$ and $C(w) = 1$ for $w = 1, 2, \ldots, N-1$, and $f(x, y)$ is the spatial distribution of the original image with $M \times N$ pixel resolution, which also represents the image of the target scene as described in the latter part of this paper; $F(u, v)$ is the frequency spectrum of the transformation where u and v are the horizontal and vertical spatial frequencies, respectively. The reconstructed image $I(x, y)$ is obtained by applying the inverse DCT.

We simplify Equations (1) and (2) into the two equations below in order to describe our technique more effectively:

$$F(u,v) = \mathrm{DCT} \cdot f(x,y) \tag{3}$$

$$I(x,y) = \mathrm{DCT} \cdot F(u,v) \tag{4}$$

where DCT represents the DCT and IDCT is the inverse transform. One of the key steps of our technique is to realize the IDCT physically to obtain the spectral components. However, the DCT in the above equation is a 2D orthogonal product of two cosine functions and contains a considerable amount of negative numbers. Therefore, we add a constant term a to both sides, equal to the average intensity of the image. This makes all elements natural numbers, as shown by Equations (5) and (6), where b represents the contrast. Then, the DCT coefficients of the object can be obtained by Equation (7). The background illumination e that exists in the experiment is added to Equations (5) and (6) as a constant, so the constant a and the background illumination value b are effectively cancelled out. We thus have

$$F_1(u,v) = (a + b \cdot \mathrm{DCT}) \cdot f(x,y) + e \tag{5}$$

$$F_2(u,v) = (a - b \cdot \mathrm{DCT}) \cdot f(x,y) + e \tag{6}$$

$$F(u,v) = \big(F_1(u,v) - F_2(u,v)\big)/(2 \cdot b) \tag{7}$$

The multiplier of $f(x, y)$ in Equations (5) and (6) can be expressed as in the Equations (8) and (9) below, respectively:

$$P_1(u,v) = a + b\frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \cos\frac{(2x+1)u\pi}{2M} \cos\frac{(2y+1)v\pi}{2N} \tag{8}$$

$$P_2(u,v) = a - b\frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \cos\frac{(2x+1)u\pi}{2M} \cos\frac{(2y+1)v\pi}{2N} \tag{9}$$

where $P_1(u, v)$ and $P_2(u, v)$ are two patterns with the spatial frequencies (u, v). Figure 1(a) and (b) show part of the two sets of the illumination patterns generated from Equations (8) and (9), respectively. The patterns are a combination of horizontal and vertical frequencies for a 64×64 ($M = N = 64$) 2D DCT. Each step from left to right and top to bottom is an increase in frequency of 1/2 cycle. For example, moving right (or down) one step from the top-left square yields a half-cycle increase in the horizontal (or vertical) frequency. The source data (64×64) are transformed to a linear combination of these

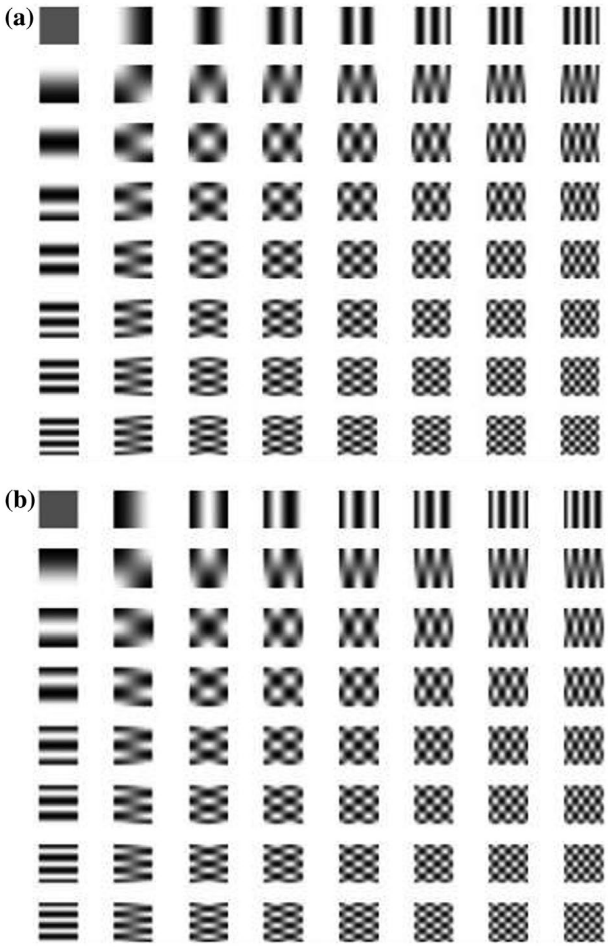

Figure 1. Part of the two sets of illumination patterns for $M = N = 64$ used in this paper. (a) Patterns generated from Equation (5), (b) Patterns generated from Equation (6).

64 horizontal and vertical frequencies. By employing the inverse 2D DCT algorithm of Equation (4), the original image can be reconstructed by the computer.

The imaging process of our proposed system can be divided into five modules, as shown in Figure 2. A projector directs the patterns generated by a computer onto the target scene, and the single-pixel detector measures the reflected light from the target scene in synchronization and transfers it to an image reconstructor. The two sets of field intensity measurements at the detector, $D_1(u, v)$ and $D_2(u, v)$, both have spatial frequencies of (u, v). Since the multiplication of $F(u, v)$ by a constant makes no difference to the reconstruction of the image by the inverse DCT, $F(u, v)$ can be calculated from Equation (10). The image of the target scene can then be reconstructed from Equation (2) by substituting.

$$F(u, v) = D_1(u, v) - D_2(u, v) \tag{10}$$

Since DCT has a strong 'energy compaction' property (*19,21*), most of the signal information tends to be concentrated in a few low-frequency components, allowing the small high-frequency components to be discarded. Thus, the technique proposed here is somewhat similar to a compressive sampling approach in that the image can be reconstructed with only a fraction of the spectra.

3. Experimental setup

The set-up for this experiment is shown in Figure 3. A digital light projector (BENQ MS513P) provides sinusoidal orthogonally structured patterns to illuminate a 3D scene consisting

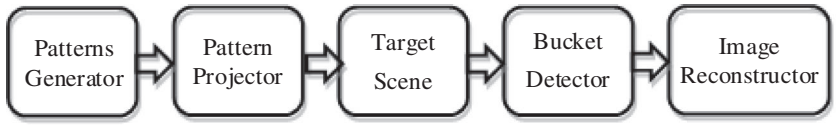

Figure 2. Procedure of computational imaging with single-pixel detectors based on 2D DCT.

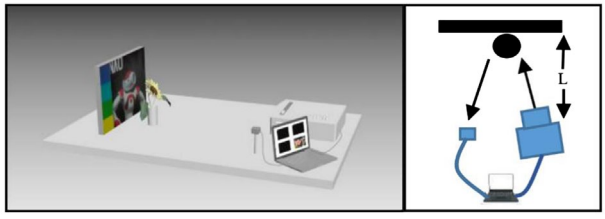

Figure 3. Experimental system for imaging a coloured 3D scene. The projector illuminates the scene with two sets of orthogonal patterns. The distance L between the projector and target scene is about 120 cm.

of a flower in front of a background picture 12 × 12 cm in size at a distance of about 120 cm. The main components of the projector are a digital micro-mirror array, a high voltage mercury lamp and a lens of 55 mm focal length. The light reflected from the scene is directed onto a CMOS sensor with a Bayer filter (Basler acA1920-155uc) which detects red (R), green (G) and blue (B) with different filters at different pixels. The signals from all the colour channels are recorded in the computer, and the full-colour image is reconstructed by the algorithm shown above.

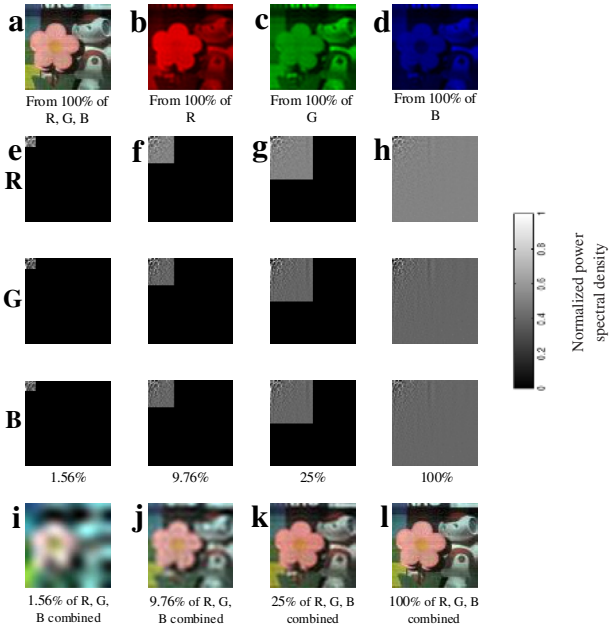

Figure 4. (a) Full-colour reconstructed image of the scene, obtained by combining the separate reconstructions from the (b) red, (c) green and (d) blue spectral data. (e)–(h) are the density distributions of the cosine transform spectra acquired for the three colours with different spectral coverages of 1.56, 9.76, 25 and 100%, with each row corresponding to red, green and blue, respectively. (i)–(l) are the reconstructed combined R, G, B images with the same spectral coverages of (e)–(h), respectively.

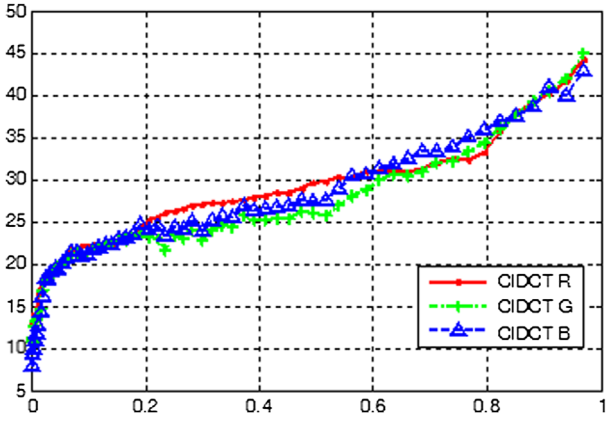

Figure 5. PSNR curves of the three different image reconstruction methods CIDCT R, CIDCT G and CIDCT B versus spectral coverage.

To synchronize the system, we designed the data collecting program written in C++ language using Microsoft Visual Studio, so that we could project the patterns and record the photodetector signals synchronously. The exposure time of each image was 0.1 s, and the patterns were switched at eight frames per second.

4. Image reconstruction

Using two groups of orthogonal sinusoidal patterns, the scene's illumination field was changed a total of $64 \times 64 \times 2$ times, for the case of $M = N = 64$. The integrated intensities of the reflected field were measured $64 \times 64 \times 2$ times, accordingly. The spectrum of each image can be calculated quickly by Equation (8). The size of the illumination patterns and final reconstruction was 64×64 pixels. The red, green and blue images of the scene are reconstructed by the algorithm based on the 2D DCT, as described above. The entire full-colour image with all the colours combined is shown in Figure 4(a). Figure 4(b)–(d) show the red, green and blue reconstructions, respectively. It should be noted that the results retrieved from any one of the three colours are graded in tone.

Figure 3(e)–(h) shows the density distributions of the cosine transform spectra for the three colours acquired with different spectral coverages of 1.56, 9.76, 25 and 100%, respectively. Since the scene image is sparse in the frequency domain, the image energy spectra, as we can see, is concentrated in the top left-hand corner of the full spectra, where the low-frequency components are located. Figure 3(i)–(l) are the reconstructed images with the same spectral coverage of red, green and blue corresponding to Figure 3(e)–(h). It is evident that the quality improves as the covering spectrum increases. The image is already recognizable when the coverage reaches 9.76%, which means that an image with 64×64 pixel resolution can be retrieved from 400×2 measurements by our technique.

For a quantitative comparison of the image quality, we introduce the peak signal-to-noise ratio (PSNR) to estimate the performance of our 2D computational imaging DCT (CIDCT) method:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[U_0(i,j) - U(i,j)\right]^2}\right), \tag{11}$$

where U_0 is the original image of the target scene consisting of $M \times N$ pixels, U is the retrieved image and $\mathrm{MAX}_I = 255$ is the maximum pixel value of the image. Naturally, the larger the PSNR value, the better the quality of the image recovered. The PSNR values of the red, green and blue components (marked as CIDCT R, CIDCT G and CIDCT B, respectively) for the different sampling coverages of the spectra in Figure 4(e)–(h) are shown in Figure 5. The PSNRs of the three components are 22.27, 21.42 and 21.29 dB, respectively, at a sampling coverage of 9.76%, which are acceptable values for recognizing the objects in the scene. When the sampling coverage exceeds 61.04%, our method could obtain high-quality images of the scene with the PSNRs larger than 30 dB.

We compare our CIDCT method with the following three methods: differential ghost imaging (DGI) (*6*), single-pixel imaging via compressive sampling (SICS) (*15*) and Fourier spectrum single-pixel imaging (FSSI) (*18*).

Table 1. PSNR values (in dB) of the different methods.

Sampling ratio (%)	CIDCT	DGI	SICS	FSSI
10	20.37	11.18	15.68	18.65
30	22.70	12.21	16.32	21.94
50	26.23	12.76	19.02	22.62

Table 2. Programme run time (in sec) of the different methods.

Sampling ratio (%)	CIDCT	DGI	SICS	FSSI
10	0.113	0.365	0.451	0.102
30	0.125	0.439	0.775	0.113
50	0.133	0.589	1.168	0.117

In SICS, we choose TVAL3 (*22*) as the reconstruction algorithm and the scrambled Walsh–Hadamard matrix that fits TVAL3 as the measurement matrix. The PSNR values for the four reconstruction methods at corresponding sampling ratios of 10, 30 and 50% are shown in Table 1. It can be seen that the image quality retrieved by CIDCT and FSSI are much better than those retrieved by DGI and SICS. The images recovered by CIDCT have slightly better quality than those recovered by FSSI in terms of PSNR values. The lengths of time required to run the reconstruction programmes are also compared in Table 2. The reconstruction algorithms of CIDCT and FSSI used here are programmed on the basis of the corresponding functions in MATLAB. CIDCT and FSSI take much less time than DGI and SICS, although the run time for CIDCT is slightly longer than that of FSSI. The run time of DGI is shorter than that of SICS, while the image quality of SICS is better than that of DGI, but both are inferior to CIDCT. Thus, we can say that our method is capable of achieving high-quality single-pixel imaging within a short reconstruction time.

5. Summary

In conclusion, this paper demonstrates a full-colour computational imaging technique based on 2D DCT with single-pixel detectors, which can be used to produce a full-colour image of a 3D scene. Compared with conventional, computational or compressive sensing GI, this method retrieves an image directly from the total light intensity of the illuminated scene, instead of having to correlate an object beam and reference beam. The number of measurements can be greatly reduced, and the image quality is very good, while the experimental set-up and data processing are simple and low cost. Furthermore, with this method, it should be possible to retrieve an image even when the light to be collected is distorted by a scattering medium. Future applications could include remote sensing using a laser modulated by a spatial light modulator to illuminate the target scene.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Natural Science Foundation of China [grant number 61473022], [grant number 60907031].

References

(1) Pittman, T.B.; Shih, Y.H.; Strekalov, D.V.; Sergienko, A.V. Optical Imaging by Means of Two-photon Quantum Entanglement. *Phys. Rev. A* **1995**, *52,* R3429–R3432.
(2) Valencia, A.; Scarcelli, G.; D'Angelo, M.; Shih, Y. Two-photon Imaging with Thermal Light. *Phys. Rev. Lett.* **2005**, *94,* 063601.
(3) Ferri, F.; Magatti, D.; Gatti, A.; Bache, M.; Brambilla, E.; Lugiato, L.A. High-resolution Ghost Image and Ghost Diffraction Experiments with Thermal Light. *Phys. Rev. Lett.* **2005**, *94*, 183602.
(4) Gatti, A.; Brambilla, E.; Bache, M.; Lugiato, L.A. Ghost Imaging with Thermal Light: Comparing Entanglement and Classical Correlation. *Phys. Rev. Lett.* **2004**, *93*, 093602.
(5) Zhang, M.; Wei, Q.; Shen, X.; Liu, Y.; Liu, H.; Cheng, J.; Han, S. Lensless Fourier-transform Ghost Imaging with Classical Incoherent Light. *Phys. Rev. A* **2007**, *75,* 021803.
(6) Ferri, F.; Magatti, D.; Lugiato, L.A.; Gatti, A. Differential Ghost Imaging. *Phys. Rev. Lett.* **2010**, *104,* 253603.
(7) Wu, J.; Liu, W.; Liu, Z.; Liu, S. Correlated-imaging-based Chosen Plaintext Attack on General Cryptosystems Composed of Linear Canonical Transforms and Phase Encodings. *Opt. Commun.* **2015**, *338,* 164–167.
(8) Zhang, D.; Zhai, Y.H.; Wu, L.A.; Chen, X.H. Correlated Two-photon Imaging with True Thermal Light. *Opt. Lett.* **2005**, *30,* 2354–2356.
(9) Scarcelli, G.; Berardi, V.; Shih, Y. Can Two-photon Correlation of Chaotic Light be Considered as Correlation of Intensity Fluctuations? *Phys. Rev. Lett.* **2006**, *96,* 063602.
(10) Chen, X.H.; Liu, Q.; Chen, K.H.; Wu, L.A. Lensless Ghost Imaging with True Thermal Light. *Opt. Lett.* **2009**, *34,* 695–697.
(11) Shapiro, J.H. Computational Ghost Imaging. *Phys. Rev. A* **2008**, *78,* 061802.

(12) Bromberg, Y.; Katz, O.; Silberberg, Y. Ghost Imaging with a Single Detector. *Phys. Rev. A* **2009**, *79,* 053840.
(13) Welsh, S.S.; Edgar, M.P.; Bowman, R.; Jonathan, P.; Sun, B.; Padgett, M.J. Fast Full-color Computational Imaging with Single-pixel Detectors. *Opt. Express* **2013**, *21,* 23068–23074.
(14) Candè, E.J.; Wakin, M.B. An Introduction to Compressive Sampling. *IEEE Signal Process Mag.* **2008**, *25,* 21–30.
(15) Duarte, M.F.; Davenport, M.A.; Takhar, D.; Laska, J.N.; Sun, T.K.; Kelly, E.; Baraniuk, R.G. Single-pixel Imaging via Compressive Sampling. *IEEE Signal Process Mag.* **2008**, *25*, 83–91.
(16) Romberg, J. Imaging via Compressive Sampling [Introduction to Compressive Sampling and Recovery via Convex Programming]. *IEEE Signal Process Mag.* **2008**, *25,* 14–20.
(17) Aβmann, M.; Bayer, M. Compressive Adaptive Computational Ghost Imaging. *Sci. Rep.* **2013**, *3,* 1545.
(18) Zhang, Z.; Ma, X.; Zhong, J. Single-pixel imaging by means of Fourier spectrum acquisition. *Nat. Commun.* **2015**, *6,* 6225.
(19) Rao, K.R.; Yip, P. *Discrete Cosine Transform: Algorithms, Advantages, Applications*, Academic: Boston, MA, **1990**.
(20) Ahmed, N.; Natarajan, T.; Rao, K.R. *IEEE Trans. Comput.* **1974**, *100,* 90–93.
(21) Makhoul, J. A Fast Cosine Transform in One and Two Dimensions. *IEEE Trans. Acoust. Speech Signal Process.* **1980**, *28,* 27–34.
(22) Li, C.B. An Efficient Algorithm for Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing. Master's Thesis, Rice University, Houston: TX, USA, **2010**.