Single-pixel camera

Tom A. Kuusela*

Department of Physics and Astronomy,

University of Turku, 20014 Turku, Finland

(Dated: July 15, 2019)

Abstract

A single-pixel camera is an interesting alternative to modern digital cameras featuring millions of pixels. A single-pixel camera is a method that produces images by exploring the object features with a series of spatially resolved patterns of light field while measuring the correlated intensity on a single detector. Nowadays single-pixel cameras are used on those applications where multi-pixel detectors are not available because the wavelength is not in visible range or light intensity is extremely low.

The spatial light modulator is an essential part of any single-pixel camera systems. They are, unfortunately, very expensive. We describe a low-cost version of single pixel camera that can be used in undergraduate physics laboratories. We show that with this camera setup students can easily demonstrate basic characteristics of computational ghost imaging and traditional raster and basis scan. Finally we explain how to perform compressive sampling of images where number of measurements is well below the actual pixel number. Compressive sampling is rapidly expanding method to perform image or signal reconstructions in many field of research.

I. INTRODUCTION

Modern digital cameras have image sensors featuring millions of pixels, and the number of pixels is the most important character for the camera performance. However, it is possible to construct a camera that has only one pixel. The idea behind this method is actually quite simple. Let us consider the normal case where we have a point source of light, which illuminates the object. Reflected light from the object is then focused on the pixel matrix of the camera. By measuring the signal from each pixel we can construct the image of the original object. But we can invert the geometrical situation in the manner of dual photography.[1] Let us put the light source in the place of the camera and collect the total light intensity into one pixel positioned in the place of the original light source. Obviously the single-pixel signal depends on the features of the object but also the structure of the light source. It is also evident that with only one measurement we cannot construct the image but many different light structures are needed. In other words, a single-pixel camera produces images by exploring the object with a series of spatially resolved light field patterns made by a spatial light modulator and measuring the correlating light intensity on the detector without any spatial resolution.

What are the most important features of single-pixel cameras?[2] Silicon-based multi-pixel sensors elements can be used only in the waveband of visible light and near infrared but not in long-wavelength infrared or deep ultaviolet where pixelated sensors are either highly expensive or they do not exist at all. In some applications a single-pixel camera can be cheaper than its multi-pixel counterpart, or it can be significantly more sensitive when using a single photon detector. Also other important features as better detection efficiency or faster temporal response could be realized with single pixel cameras. But on the other hand, single-pixel cameras need more numerical processing power, which can be a limiting factor. During the last decade single-pixel cameras have demonstrated applications in many fields, for example in multispectral imaging,[3] real-time 2D and 3D video,[4,5], time-of- flight measurements[6] microscopy,[7] holography[8] and X-ray tomography[9].

In the following we first describe a low-cost version of the single-pixel camera system suitable for a student exercise laboratory. Next we demonstrate various scanning methods based on the concept of ghost imaging, raster scan, basis scan and finally compressive sampling.

II. CAMERA CONSTRUCTION

As described in the previous chapter, the spatial light modulator (SLM) is the main component of the single-pixel camera. There are two types of SLMs: digital micromirror devices (DMD) and liquid-crystal devices (LCD). DMDs are used in digital light projectors to display films and presentation slides, and most of them are based on Texas Instrument's components. The DMD consists of an array of electrostatically actuated micromirrors, each of having the side length of about 10μm, and they can be individually tilted $\pm 10^\circ$. There are some cheap developer kits of TI DMDs but the control electronics and software are rather complex and they need lot of expertise in construction. Transmissive and especially reflective LCDs are excellent research tools in control of amplitude and phase of light fields but they are, unfortunately, too expensive for student exercise laboratories.

We have designed a simple low-cost version of SLM (see Fig. 1). The patterns of structured light are displayed on the computer screen by a simple software. We used 50 x 50 pixel matrix (the actual size 13 x 13 mm^2 on the screen, the pixel size 0.265 mm), and in most our experiments each pixel is either black or white corresponding to the values 0 and 1. The object is a mask which is a pattern printed on the transparent sheet and positioned almost in contact to the screen (distance less than 1 mm) in front of the pixel matrix. As a detector we used a large-area photodiode (OSD100-E by Centronic). As the diameter of the active area of this photodiode is about 11 mm, almost all light can be captured without any collecting optics if the distance from the screen is within few millimeters. In all our measurements the distance of the detector from the mask was 2 mm. It is also possible to use a smaller photodiode with a large-diameter positive lens of short focal length.

The photodiode signal is amplified to the level suitable for the analog-digital converter (Picolog 1216 12-bit datalogger by Pico technology, sampling rate 50 kHz). Typically modulation in the photodiode signal is very low and thus high amplification is needed. In the computational ghost imaging and basis scan (these are explained in the following chapter) on average half of the pixels are illuminated, therefore the actual important part of the signal is on the top of a large background and amplified output signal can be too large for the dynamical range of the analog-digital converter. This background can be eliminated (or reduced to an adequate level) in the output by adjusting (potentiometer P1) the offset voltage in the second amplification stage. This adjustment can help to eliminate also the

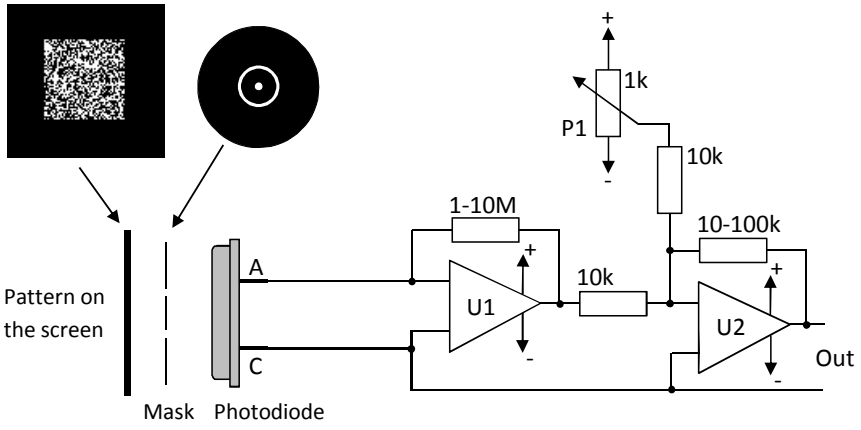

FIG. 1. Experimental setup: random pixel patterns are shown on the computer screen. The light intensity of the pattern trough the mask is measured by a large area photodiode. Photodiode current is amplified with two operational amplifier stages. The potentiometer P1 sets the voltage offset on the output signal.

the effect of ambient light.

Even when the pixel matrix is not changing the signal from the photodiode is a train of short pulses because of the refreshing cycle of the display. With our monitor the pulse frequency was 240 Hz and the pulse width 2.5 ms. For each pattern we captured 30 ms of the signal (totally 7-8 pulses) and simply searched the maximum value, which corresponds well to the light intensity. By searching the amplitude of each pulse and averaging them we can produce slightly better results.

A similar type of lensless single-pixel imaging system is demonstrated as a compact scanner which has no moving parts.[10]. In that setup the paper document is tightly sandwiched in a LCD and solar cell detector, and the total thickness of the scanner is less than 3 mm. This scanner is, unfortunately, very slow because the rise and fall time of the solar cell are rather long and the LCD controller does not provide any hardware synchronization, thus a complete scanning takes several hours.

III. SCANNING METHODOLOGIES

A. Computational ghost imaging

Ghost imaging (GI) is an old, perhaps the first, variant of single-pixel cameras in the modern form, and especially a certain version of GI is closely related on the current topic. In the original classical ghost imaging method a (pseudo)thermal light beam is divided into two identical part, the reference and object arm, by a beam splitter.[11,12] In the reference arm the light field $I_i(x, y)$ is captured by a CCD camera (or another device having spatial resolution). At the object arm a bucket detector (actually a single-pixel detector) measures the total intensity s_i, which is transmitted through the object with the transmission function $M(x, y)$:

$$s_i = \int dxdy I_i(x, y) M(x, y). \tag{1}$$

In order to reconstruct the transmission function of the object, the total intensities are cross-correlated with the intensities measured in the reference and averaged over several realizations. If the light field is spatially incoherent, we obtain

$$M_{GI}(x, y) = \langle s_i \Delta I_i(x, y) \rangle_N = \frac{1}{N} \sum_{i=1}^{N} s_i \Delta I_i(x, y), \tag{2}$$

where $\Delta I_i(x, y) = I_i(x, y) - \langle I_i(x, y) \rangle_N$. By subtracting the average $\langle I_i(x, y) \rangle_N$ we can avoid a noise offset term, which would otherwise reduce the visibility of the ghost image. We can see that the ghost image measurement is a vector projection of the transmission function over N different random functions.

In this image construction scheme the physical thermal source can be replaced with a light source, which is *deterministically* controlled by a spatial light modulator. Since the pattern configuration of SLM is now *known* it is not any more necessary to divide the light beam into the object and reference arm. If the light field is incoherent the result (2) is still valid. This image construction method is called *computational ghost imaging* (CGI)[13].

In our camera system we constructed random pixel patterns by generating pseudo-random numbers in the range $[0, 1]$. If the random number ≤ 0.5 the pixel was black, otherwise white, thus the probability was equal for these two choices. The object (see Fig. 1) was a simple ring (the outer diameter 7.0 mm, width 0.8 mm) with a central dot (the diameter 1.2 mm). The ring and the dot were fully transparent and other parts as opaque as it was possible

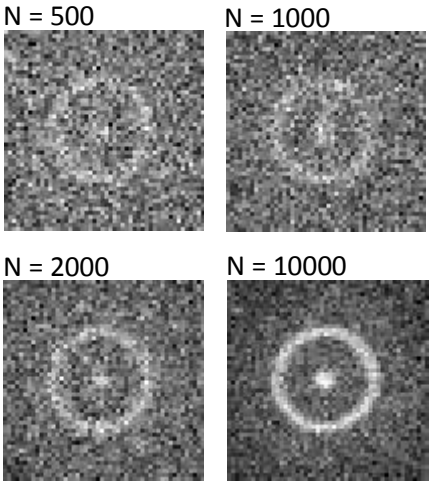

FIG. 2. Computational ghost imaging reconstructions made by a single pixel camera using random pixel patterns with $N = 500$, 1000, 2000 and 10000 realizations.

with the laser printer. Single pixel camera reconstructed images with various numbers of realizations are shown in Fig. 2. We can see that with 500 realizations a noisy shadow of the original image is visible but low noise images need 10000 realizations or even more.

B. Raster and basis scan

By CGI it is possible to reconstruct the image but high number of realizations must be created if a high signal-to-noise ratio is needed. The most simple way to capture the image is use *raster scanning*: the object is illuminated pixel-by-pixel. This method is commonly used in applications where multi-pixel detectors are not available. Modern scanning system consist of a pair of galvanometer mirrors that are used to steer either illumination or detection onto a single-pixel detector. In our camera system we need 2500 different pixel patterns, each having only one white pixel (and all the rest black). The result of the raster scan is shown in Fig. 3 (top left). The original image is quite well reconstructed but it is clearly smoothed obviously since the pixels of the screen radiate some light also in other directions than just forward depending on the angular intensity distribution. Even when the object is close the screen some pixels outside the transparent areas might be slightly seen by the detector. There are also some periodical patterns due to interference of the screen pixels and the pixel structure of the laser printed shape, even the totally black areas. Because of this interference more complex non-binary picture masks are not properly reconstructed

since with laser printers gray levels are actually produced by raster patterns.

Let us consider the 2-D array of N pixels as an $N \times 1$ column vector with elements x_n, $n = 1, 2, \ldots, N$, i.e., we treat an image by vectorizing it into a long one-dimensional vector. Any vector in $\mathbb{R}^N$ can be represent in the terms of a basis of (orto)normal $N \times 1$ vectors $\{\psi_i\}_{i=1}^N$. Forming the coefficient vector s and the $N \times N$ basis matrix $\Psi = [\psi_1|\psi_2|\ldots|\psi_N]$ by stacking the vectors $\{\psi_i\}$ as columns, we can write the image in the form

$$x = \sum_{i=1}^{N} s_i \psi_i \tag{3}$$

or $x = \Psi s$. Weighting coefficients are determined by the inner product $s_i = (x, \psi_i) = \psi^T x$ (T denotes transposition). Our single pixel camera just performs, at least approximately, this "computation". As an example, in the raster scan method each vector ψ_i has only one non-zero element. This is a special case of the basis. Also Eqs. (1) and (2) of CGI can be written in a similar fashion.

There are other, sometimes more useful, sets of basis functions. Often we want to find a basis where the coefficient vector s is *sparse* (it has only $K \ll N$ non-zero coefficients) or *compressible* (there are just few large coefficients and many small ones). Natural images tend to be compressible in the discrete cosine transform (DCT),[14] which is the core of the JPEG compression algorithm.[15] There are several slightly different versions of DCT but we use the so-called DCT-II:

$$X_k = \sum_{n=1}^{N} x_n \cos\left[\frac{\pi}{N}\left((n-1) + \frac{1}{2}\right)(k-1)\right], k = 1, \ldots, N. \tag{4}$$

The basis vectors for Ψ are easily calculated from the transform (4) by setting $x_1 = (1, 0, \ldots, 0)$, $x_2 = (0, 1, \ldots, 0)$ and so on. We scaled the component value range of $[-1, 1]$ of the basis vectors to the gray-scale of the pixels (not just black and white). The result of the image construction with these 2500 basis vectors, based on the equation (3), is shown in Fig. 3 (top right). The quality of the image is almost the same as in the raster scan method, as expected, but the pixel interference is somewhat more prominent. The coefficients s_i, magnitude sorted from the largest value to the smallest, as a function of the vector component index are shown in Fig. 4. We can see that there are relatively few large coefficients and they all decay approximately under the power law, typical for compressible signals, thus the image is compressible but not very efficiently: 500-1000 basis vectors are needed.

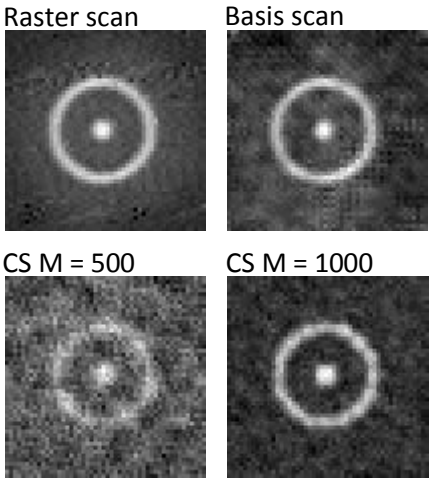

FIG. 3. Image reconstructions made by the raster scan (top left), discrete cosine transform (DCT) basis scan (top right) and compressive sampling (CS) method (bottom left and right).

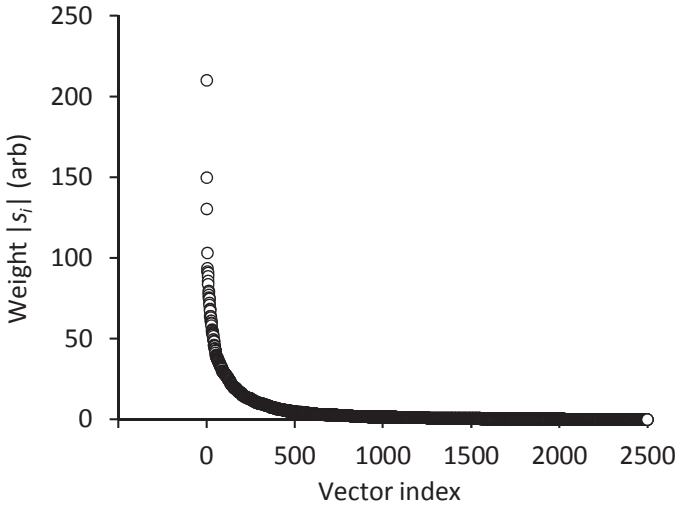

FIG. 4. The magnitude sorted weight coefficients $|s_i|$ as a function of the vector component index.

C. Compressive sampling

When using DCT or other similar type of bases it is possible to construct a compressed form of the image. But there is an inherent disadvantage: we must always perform N measurements and *afterwards* select those K basis vectors which are important, even if the desired K is very small. *Compressive sampling* (CS) method bypasses the sampling process

and directly acquires a compressed reconstruction using only $M < N$ measurements between the image x and test vectors $\{\phi_m\}_{m=1}^{M}$.[16–18] As a result we get inner products $y_m = (x, \phi_m)$. Measurements y_m can be arranged into a $M \times 1$ vector y and the test vectors ϕ_m as rows into a $M \times N$ matrix Φ. By substituting $x = \Psi s$ from Eq. (3) we get

$$y = \Phi x = \Phi\Psi s = \Theta s, \tag{5}$$

where Θ is $M \times N$ matrix. We assume that Φ does not depend in any way on the image x, thus the measurement process is totally nonadaptive. The transformation from x to y is a dimensionality reduction in which some information is lost. From a mathematical point of view the main problem is to find a matrix Φ such that information in any sparse or compressible image (or in a signal in general) is not corrupted by this reduction. Amazingly this is possible: it can be shown that with high probability random matrices are good choices.[19] Intuitively this procedure can be understood in the following manner: each random matrix collects a small amount of information over all frequencies (or other features depending on the compressive basis Ψ) if there is no "interference" between Φ and Ψ. This property can be formulated as a rigid mathematical condition called *restricted isometry property*.[16]

We have N-dimensional vector x and we take only $M < N$ measurements, the vector y. Since $M < N$, the problem is ill-conditioned. If, however, x is sparse and there are K non-zero coefficients in s, this problem can be solved if $M \geq K$.[20,21] The problem can be formulated as a minimization of the following functional of s

$$\|\Theta s - y\|^2 + \lambda \|s\|_p, \tag{6}$$

where the regularization parameter λ controls the trade-off between the data misfit $\Theta s - y$ and the penalization in the term of p -norm of the parameters. The parameter λ is closely related to the noise level of the data y. If $p = 2$ ($\|s\|_2 = \sqrt{\sum_i s_i^2}$) we have the well-known method of least squares, which as a linear problem can be easily solved. Unfortunately, the penalty term of this type is generic and does not utilize any *a priori* information about the desired solution s, and it can hardly find any sparse solution. The case of $p = 0$ (counts the number of non-zero components in s) would measure directly the sparsity of the solution but this problem is numerically unstable and very difficult to solve.

The choice $p = 1$ ($\|s\|_1 = \sum_i |s_i|$) can, surprisingly, exactly recover K-sparse signals and well approximate compressible ones when $M \geq K \log(N/K)$.[20,21] This optimization problem

often called "basis pursuit" can be solved using standard convex programming algorithms. We will not go to the details of these algorithms since they are quite complex and we did not try to code it by ourselves. Instead, we used freely available toolbox *L1Packv2* designed for Mathematica.[22] There are other similar free packages (*Sparse-Lab*, l_1-*magic*) for Matlab. See Appendix for the details of *L1Packv2*.

We made experiments with the same original image as earlier. The matrix Ψ consists of the same DCT basis vectors as used in the basis scan method. The test functions ϕ_m are those random matrices used in CGI method. Results of compressive sampling with $M = 500$ and $M = 1000$ are shown in Fig. 3 (bottom left and right). Clearly 500 samples is not enough for proper reconstruction of the image but with 1000 samples we achieve almost the same quality as in the raster and basis scan. If we compare the CS image with the CGI image ($N = 1000$) we find significant increasing in the signal-to-noise ratio (SNR).[23] We obtain $SNR_{CGI} = 1.5$ and $SNR_{CS} = 14.9$. It should be noted that in the measurement phase we collect exactly the same information in CGI and CS methods but the last one needs much more computational efforts.

Compressive sampling has several advances. First, CS method can be called universal since the same random matrix Φ can be used with many different bases Ψ. In CS method each measurement collects the same amount of information and the lost of few measurements does not corrupt the reconstruction, in contrast to the basis scan where collected data is strictly hierarchical. This feature of CS enables also progressive data collection: we obtain better reconstruction as we get more measurements.

IV. CONCLUSION

Single pixel cameras are nice examples on the change of the viewpoint, even literally. We describe a simple single-pixel camera setup where the expensive spatial light modulator is replaced by pixel patterns on the computer monitor screen. Although there are limitations in the optical setup all essential features can be demonstrated. Students can prepare own objects with a laser printer and explore the concept of ghost imaging and make experiments with various sparsitfying or compressing bases in image construction. It is possible that students also design own bases and check their performance. Since the data collection is rather slow, because of low refreshing rate of the screen, image reconstruction takes time but

this can be converted to an advantage: reconstruction can be followed in "real" time. The compressive sampling is an advanced topic but if there is any access to tools like Mathematica or Matlab, students learn more on this denoising method whose use is expanding rapidly in many fields of physics.

Appendix A: Use of Mathematic l_1-norm minimizer package L1Packv2

In our software the measurement vector y was written into a file as a single column. This data can be read into Mathematica by the command

y = Import["YFileName.txt", "List"]

The Θ matrix was written into a file as a comma separated columns. This data can be read into Mathematica by the command

Theta = Import["ThetaFileName.csv"]

The main minimizing algorithm in L1Packv2 is *FindMinimizer*. The user should set a stop condition for this algorithm. There are two most suitable for our purpose: *MinimumDiscrepancy* $\rightarrow d$ and *MaximumNonZero* $\rightarrow M$. With the first one the calculation will be stopped when $\|Thetas - y\|^2$ has reached the value d. Normally d is the square of the noise level of the signal. With the second condition the algorithm will stop when there are M non-zero component in the minimizer s. In our examples we used this condition and the corresponding command in Mathematica is

s = FindMinimizer[Theta,y,MaximumNonZero → 500 or 1000]

The result vector s can be written into the file by the command

Export["SFileName.txt",s]

---

* tom.kuusela@utu.fi

[1] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch,"Dual photography," ACM Trans. Graph. **24**, 745–755 (2005).

[2] M. P. Edgar, G. M. Gibson, and M. J. Padgett,"Principles and prospects for single-pixel imaging," Nature Photon. **13**, 13–20 (2019).

[3] F. Rousset, N. Ducros, F. Peyrin, G. Valentini, C. D'Andrea, and A. Farina ,"Time-resolved

multispectral imaging based on an adaptive single-pixel camera," Opt. Express **5**, 10550–10558 (2018).

[4] 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. **5**, 10669-1–8 (2015).

[5] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3D single-pixel video," J. Opt. **18**, 035203-1–7 (2016).

[6] M.J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, "Single-pixel three-dimensional imaging with time-based depth resolution," Nat. Comm. **7**,12010-1–6 (2015)

[7] N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, "Single-pixel infrared and visible microscope," Optica **1**, 285–289 (2014).

[8] P. Clemente, V. Duran, E. Tajahuerce, P. Andres, V. Climent, and J. Lancis, "Compressive holography with a single-pixel detector," Opt. Lett. **38**, 2524–2527 (2013).

[9] J. Greenberg, K. Krishnamurthy, and D. Brady,"Compressive single-pixel snapshot x-ray diffraction imaging," Opt. Lett. **39**, 111–114 (2014).

[10] Z. Zhang, Z. Su, Q. Deng, J. Ye, J. Peng, and J. Zhong,"Lensless single-pixel imaging by using LCD: application to small-size and multi-functional scanner," Opt. Exp. **27**, 3731–3745 (2019).

[11] R. S. Bennink, S. J. Betley, and R. W. Boyd,"Two-photon coincidence imaging with a classical source," Phys. Rev. Lett. **89**, 113601-1–4 (2002).

[12] L. Basano and P. Ottonello, "Ghost imaging: Open secrets and puzzles for undergraduates," Am. J. Phys. **75**, 343–351 (2007).

[13] J. H. Shapiro, "Computational ghost imaging,"Phys. Rev. A **78**, 061802-1–4 (2008).

[14] N. Ahmed, T. Natarajan, and K. R. Rao,"Discrete cosine transform," IEEE Trans. Comp. C **23**, 90–93 (1974).

[15] G. K. Wallace,"The JPEG still picture compression standard," Commun. ACM **34**, 30–44 (1991).

[16] R. G. Baraniuk, "Compressive sensing," IEEE Signal Process. Mag. **24**, 118–120 (2007).

[17] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baranuk,"Single-pixel imaging via compressive sampling," IEEE Signal Process. Mag. **25**, 83–91 (2008).

[18] O. Katz, Y. Bromberg, and Y. Silberberg,"Compressive ghost imaging," Appl. Phys. Lett. **95**, 131110-1–4 (2009).

[19] E. J. Candes and T. Tao,"Near-optimal signal recovery from random projections: universal encoding strategies?" IEEE Trans. Inf. Theory **52**, 5406–5425 (2006).

[20] E. Candes, J. Romberg, and T. Tao,"Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,"IEEE Trans. Inform. Theory **52**, 489–509 (2006).

[21] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 1289–1306 (2006).

[22] I. Loris,"L1Packv2: A Mathematica package for minimizing an l_i-penalized functional," Comp. Phys. Comm. **179**, 895–902 (2008).

[23] Also in image analysis the signal-to-noise ratio is commonly determined as the ratio of the average signal and the standard deviation of the background noise. The signal average was calculated over those pixels where the mask was totally transparent and the background level over those pixels where the mask was opaque.