



Single-pixel imaging 12 years on: a review

GRAHAM M. GIBSON,^{1,2}  STEVEN D. JOHNSON,^{1,3}  AND MILES J. PADGETT^{1,4} 

¹*School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK*

²*graham.gibson@glasgow.ac.uk*

³*steven.johnson@glasgow.ac.uk*

⁴*miles.padgett@glasgow.ac.uk*

<https://www.gla.ac.uk/schools/physics/ourresearch/groups/optics/>

Abstract: Modern cameras typically use an array of millions of detector pixels to capture images. By contrast, single-pixel cameras use a sequence of mask patterns to filter the scene along with the corresponding measurements of the transmitted intensity which is recorded using a single-pixel detector. This review considers the development of single-pixel cameras from the seminal work of Duarte et al. up to the present state of the art. We cover the variety of hardware configurations, design of mask patterns and the associated reconstruction algorithms, many of which relate to the field of compressed sensing and, more recently, machine learning. Overall, single-pixel cameras lend themselves to imaging at non-visible wavelengths and with precise timing or depth resolution. We discuss the suitability of single-pixel cameras for different application areas, including infrared imaging and 3D situation awareness for autonomous vehicles.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

The concept of single-pixel imaging followed the development of compressive sensing [1–4] and was reported soon after in a seminal paper by Duarte et al. at Rice University [5]. This pioneering work is a combination of different imaging and sampling techniques which has inspired the field of single-pixel imaging, laying the foundations for recovering images from a single-pixel camera when the number of measurements is fewer than the total number of unknown pixels in the image, that is, when the properties of the image are sensed compressively, also known as under-sampling or sub-sampling.

Prior to this work, in 2005, Sen et al. had published the paper “Dual Photography” [6] which proposed the idea that an image could be captured using just a single photodetector (single-pixel detector) rather than a detector array as used by most common imaging devices such as mobile phones and digital SLR cameras. Here, the spatial structure is provided by interrogating a scene with a series of spatially resolved patterns while measuring the correlated intensities using the single-pixel detector. The development of silicon-based charge-coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) pixelated sensors has brought the benefits of cheap, high-performance, imaging technologies for many applications in the visible (VIS) wavelength spectrum. However, single-pixel detectors can bring significant performance advantages such as sensitivity at non-visible wavelengths or very precise timing resolution, both of which can be impractical or prohibitively costly to implement as a pixelated imaging device.

A popular choice for non-visible wavelength single-pixel imaging has been in the short-wave infrared (SWIR) spectral region (approximately 1–3 μm) due to the availability of detectors having a good sensitivity [7,8]. In particular, telecoms research has provided a range of InGaAs devices which has allowed both cost-effective detectors and illumination sources to be developed (operating in the 800 nm to 1800 nm range). This wavelength range has been shown to be particularly suited to imaging through scattering media, such as smoke [8], and has also been used to detect and image hydrocarbon gas leaks [9].

Single-pixel imaging has provided an ideal test platform for new state-of-the-art detector technologies, allowing the development of cost-effective imaging systems at wavelengths across the electromagnetic spectrum. Examples include x-ray imaging [10–12], terahertz imaging [13–15], compressive radar [16], a VIS-NIR telescope [17] and fluorescence microscopy [18]. They have also utilised various sampling schemes including compressive sensing and machine learning. Figure 1 shows a timeline of the development of a range of single-pixel imaging systems, including a range of modulation technologies and sampling schemes.

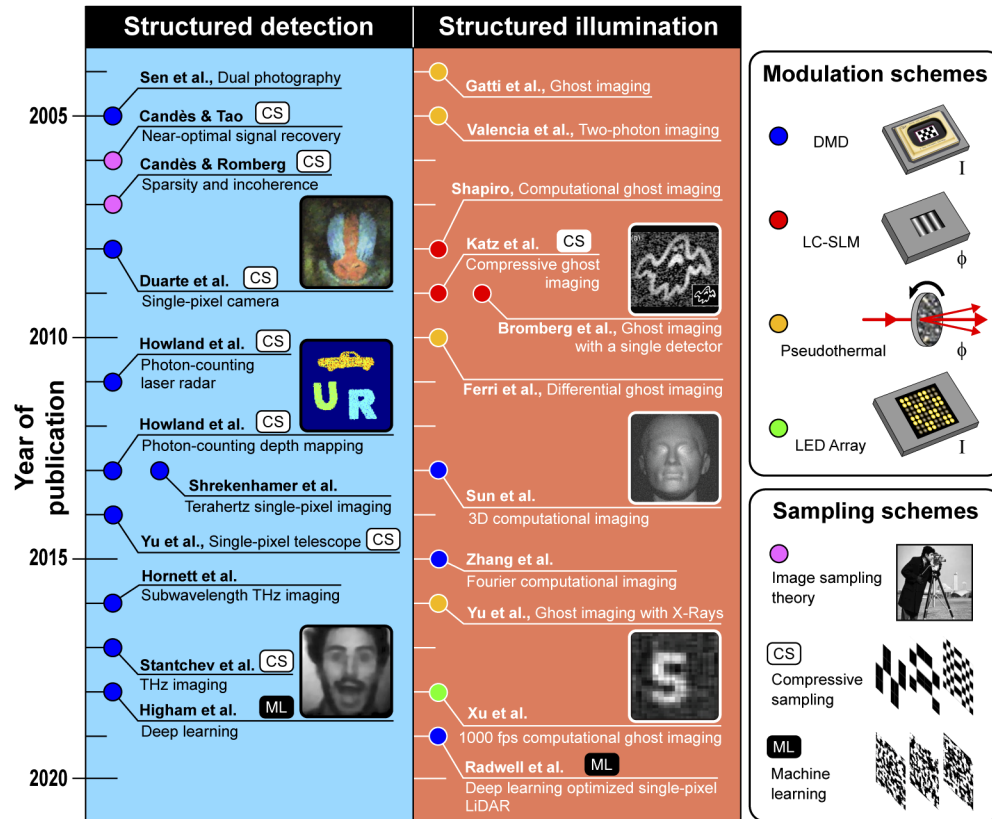


Fig. 1. Timeline of developments in single-pixel imaging. Publications are shown by year and highlight the modulation technology and sampling scheme used. It is interesting to note that systems based on a structured detection approach, and employ sampling schemes such as compressive sensing (CS) or machine learning (ML), are often termed single-pixel cameras, whereas those based on structured illumination are often referred to as computational ghost imaging. The following references are shown: Sen 2005 [6], Candès 2006 [4], Candès 2007 [19], Duarte 2008 [5], Howland 2011 [20], Howland 2013 [21], Shrekenhamer 2013 [14], Yu 2014 [17], Hornett 2016 [15], Stantchev 2017 [22], Higham 2018 [23], Gatti 2004 [24], Valencia 2005 [25], Shapiro 2008 [26], Katz 2009 [27], Bromberg 2009 [28], Ferri 2010 [29], Sun 2013 [30], Zhang 2015 [31], Yu 2016 [11], Xu 2018 [32] and Radwell 2019 [33].

A particular application that has recently attracted much attention is single-pixel imaging using time-of-flight (ToF) measurements, which can be used to recover 3D profiles of a scene from a distance. When combined with the recent advances in machine learning algorithms, single-pixel imaging shows promise as a powerful technique for low-cost, scan-free, 3D sensing and classification. This paper provides a review of single-pixel imaging techniques, including that of the closely related field of computational ghost imaging (computational GI), and focuses

on the main algorithm and hardware developments over the past twelve years. There are still many discussions on the distinction between single-pixel and computational GI. In this paper we discuss them with respect to their common terminology in the literature; single-pixel imaging often seeks to solve an inverse problem, whereas computational GI often seeks to perform a reconstruction from an ensemble average.

2. Basics of single-pixel imaging

A simple method of capturing an image using a single-pixel detector is to sequentially measure each pixel in turn, as in the raster-scan approach used in the original mechanical televisor of John Logie Baird [34]. However, sequentially measuring information on only one pixel in turn is an inefficient use of the available illumination light. A more common scan strategy is to use a sequence of spatially resolved patterns and to record the intensity measurements of the correlations between the patterns and the object, or scene. This correlation measurement can be performed in one of two ways. A light modulator placed in the image plane of a camera lens can be used to mask images of the scene, the filtered intensities being measured by the single-pixel detector. This mode of operation is commonly referred to as structured detection (see Fig. 2), and is often used in the field of single-pixel imaging or single-pixel cameras. Alternatively, the light modulator can be used to project patterns onto the scene and the single-pixel detector used to measure the back scattered intensities. This mode of operation (shown in Fig. 3) is commonly referred to as structured illumination, and is often used in the field of computational GI (discussed in detail in section 3). In both these configurations the conventional light source can be replaced with a pulsed laser (also shown in Fig. 3) so as to provide time-of-flight information and hence depth of the scene (discussed in detail in section 6.2). Section 4 discusses some of the modulation technologies commonly used in single-pixel imaging.

The object or scene can be reconstructed by multiplying each pattern in the sequence by the corresponding single-pixel intensity measurement, resulting in a set of weighted patterns that can be summed to form an image. In principle, reconstructing an image comprising of N pixels in total requires a sequence of $M = N$ different patterns. However, if the set consists of non-orthogonal patterns and / or the measurements are subjected to noise then a large number $M \gg N$ measurements are needed in order to achieve a good signal to noise ratio (SNR) of the final image. A common approach is to use an orthogonal pattern set, such as the Hadamard basis (see section 5), and measure the differential intensity for each pattern and its contrast inverse (i.e. photographic negative).

Given a sequence of N -element orthonormal pattern pairs $P_{(x,y),m}$ (where m is the pattern sequence number), the corresponding differential intensity signals between the positive and inverse patterns are S_m , which are proportional to the correlations between each pattern and the scene. Based on M patterns, the 2D image estimate of the object or scene, $O_{(x,y),M}$, can be obtained by

$$O_{(x,y),M} = \frac{1}{M} \sum_{m=1}^M S_m P_{(x,y),m} \quad (1)$$

It is clear that a means of significantly reducing the number of required patterns, and measurements, is necessary for single-pixel imaging systems to be widely adopted. Compressive sensing (CS) [3,35,36] has been shown to be a route for exploiting the redundancy in the structure of most natural signals or images. CS is based on the principle that most natural images are sparse when expressed in the appropriate basis, i.e. a basis having many coefficients that are close, or equal, to zero. This is the case for image compression algorithms such as JPEG [37,38] or JPEG 2000 [39]. CS enables image reconstruction with far fewer measurements than are required for conventional sampling schemes, allowing faster data acquisition or higher SNR [27]. However, despite the focus on faster imaging, or imaging with improved SNR , many sensing problems do

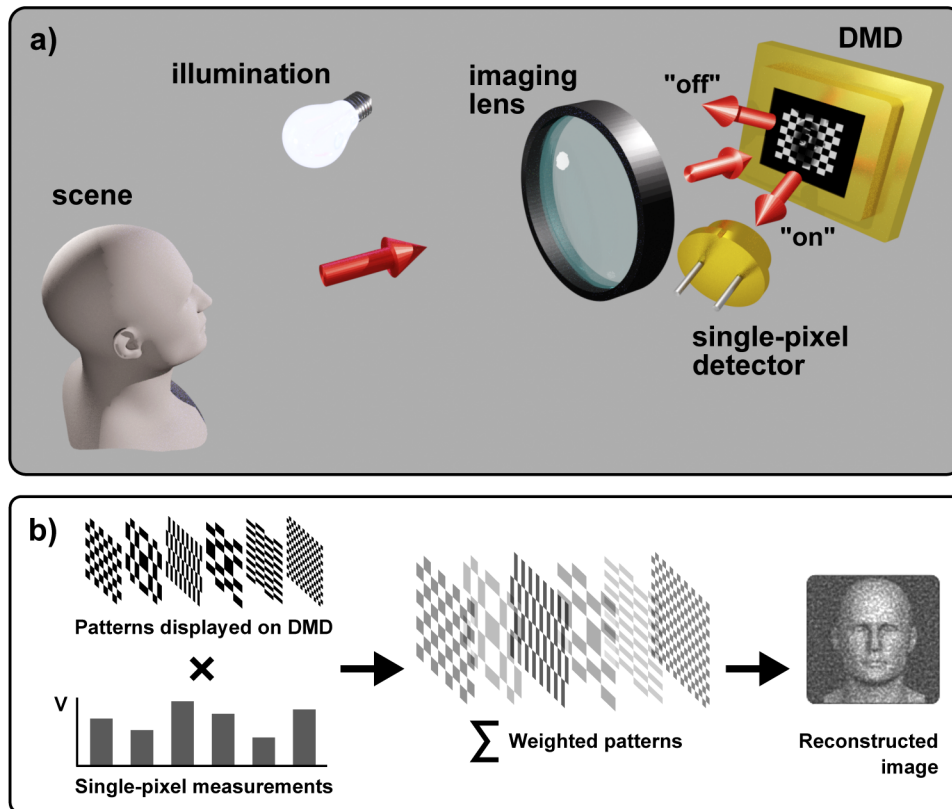


Fig. 2. Structured detection setup. a) A digital micromirror device (DMD) can be used to spatially filter light by selectively redirecting parts of an incident light beam at $\pm 24^\circ$ to the normal, corresponding to the individual DMD micromirrors being in the "on" or "off" state respectively. An object is flood-illuminated and imaged onto the DMD, where a sequence of binary patterns displayed on the DMD can be used to mask, or filter, the image. A single photodetector is used to measure the total filtered intensity for each mask pattern, allowing an image of the object to be reconstructed. b) Each pattern in the sequence is then multiplied by the corresponding single-pixel intensity measurement to give a set of weighted patterns that can be summed to reconstruct the image.

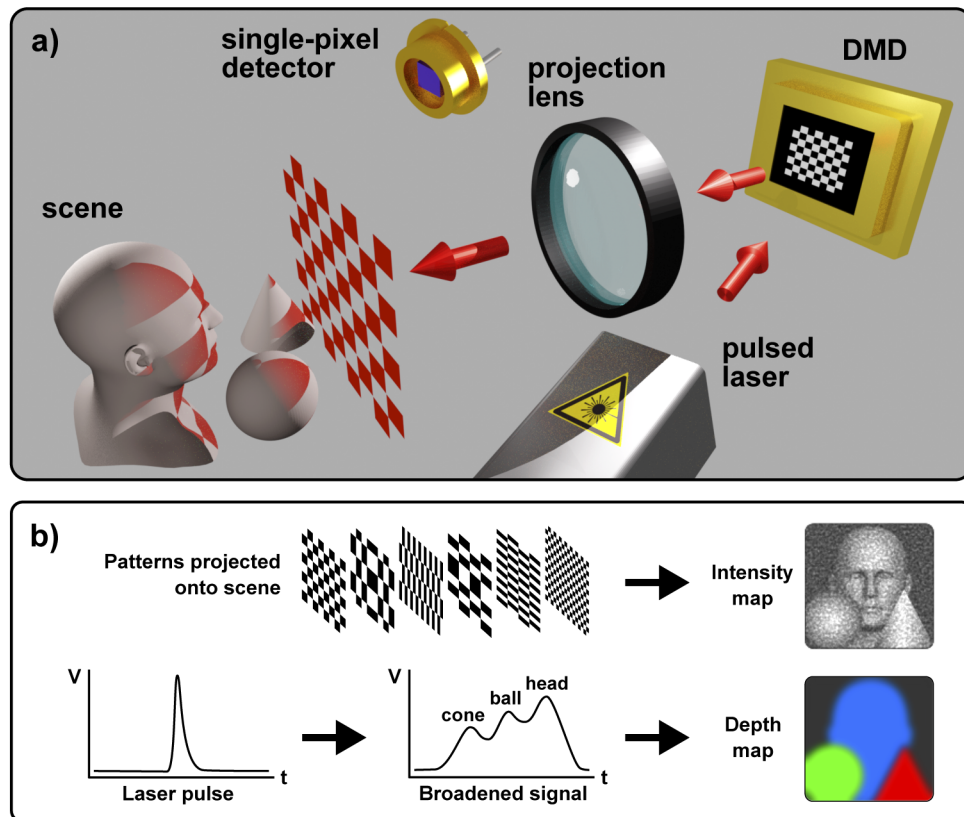


Fig. 3. Structured illumination with time-of-flight. a) In an alternative configuration, the DMD is used to project a sequence of light patterns onto a scene and the single-pixel detector measures the total back scattered intensity. For both structured illumination and structured detection a pulsed laser can be used as the illumination source to perform temporal resolution measurements using a single-pixel detector (as shown here). Recording the temporal form of the back scattered light provides a measure of the distance travelled by the light and hence depth of the scene. b) Similar to the structured detection scheme, the sequence of projected patterns and the corresponding intensity measurements allows an image to be reconstructed. In the case where a pulsed laser is used, the additional time-of-flight information from the broadened back scattered pulse allows a depth map of the scene to be constructed.

not require the full signal to be reconstructed. This is the case in applications such as detection or classification [40]. In the case of compressive classification, the resulting dimensionally reduced matched filters are sometimes termed “smashed filters” [41]. Image-free classification is also discussed in section 8 when using machine learned sampling schemes.

3. Computational ghost imaging

A field that is very closely related to single-pixel imaging is that of ghost imaging (GI), which is a technique that exploits the quantum nature of the entangled photon pairs produced in spontaneous parametric down-conversion [42]. A pump laser incident on a nonlinear crystal produces the photon pairs, often termed the *signal* and *idler*, which are entangled in their positions and hence, measuring the position of one implies the position of the other. The *signal* and *idler* beams are separated along different paths, one is measured by a spatially resolved detector such as a CCD or scanning pinhole and photodetector, the other interacts with the target object and is collected by a single-pixel detector (in GI this is often referred to as a bucket detector). Importantly, the light captured by the CCD never interacts with the target object. Only by correlating the CCD and bucket detector measurements can the “ghost” image be revealed [43]. Whilst originally demonstrated using degenerate *signal* and *idler* photons at 702 nm, GI has also been achieved at other wavelengths, including a demonstration using non-degenerate photons at 1550 nm and 460 nm [44]. GI has even been achieved using two beams formed by correlated pairs of ultracold metastable helium atoms [45].

However, it was soon realised that while GI was originally designed to exploit the quantum nature of light, it was also possible to be performed in a classical experiment [24,46]. Similar to the quantum experiment, a structured illumination light field is split into two near identical beams, usually termed the reference and object beams. The reference beam is recorded by the CCD while the object beam impinges upon the target object, and the scattered or transmitted light is then measured by the bucket detector. Bennink et al. [47] demonstrated coincidence imaging using a classical light source made by chopping and deflecting a laser beam, creating pairs of angularly correlated pulses. However, in most of the early examples of classical GI the object being imaged was illuminated by a time-dependent speckle pattern, generated by passing a collimated laser beam through a rotating ground-glass diffuser [25,48,49] (see section 4.1 for a discussion on pseudothermal modulation schemes). A simple beamsplitter copies this pseudothermal source into the reference and object beams.

The classical form of GI was developed further by Shapiro [26], who proposed the use of a computer controlled spatial light modulator (SLM) for creating the speckle patterns to illuminate the object. Since the patterns are predetermined using a computational method, the beamsplitter and CCD sensor are no longer required as it is no longer necessary to record the illumination beam, only the synchronised intensity measurements from the bucket detector are required in order to reconstruct the image. This form of GI, often referred to as computational GI, was demonstrated experimentally by Bromberg et al. [28] and shortly afterwards was demonstrated experimentally using compressive sensing [27]. Erkmen and Shapiro [50] provide a useful review of quantum, classical and computational ghost imaging.

Similar to the discussion in section 2, if the number of resolution cells “speckles” within the illumination pattern is N , one needs in principle at least $M = N$ different patterns in order to fully reconstruct the image of the object. In practice, since these correlation methods are statistical in nature, there is spatial overlap between the different speckle patterns and hence they form a non-orthogonal measurement basis. A large number $M \gg N$ measurements are therefore needed in order to achieve a $SNR \gg 1$ [27]. A major downside to classical GI was the large background level in the reconstructed images compared to that achieved using a quantum source. Methods of improving the SNR of GI systems were soon proposed, with differential GI being the most

widely adopted [29]. Here, a differential bucket detector signal measurement is employed which is sensitive only to the fluctuating part of the intensity signal.

There have been useful comparisons of computational GI systems to single-pixel cameras [51]. In particular, computational GI can be compared to the original work on dual photography [6] which is a novel photographic technique that exploits Helmholtz reciprocity to interchange the lights and cameras in a scene [52,53]. This can also be compared to the work of Sun et al. [30] where four spatially separated single-pixel detectors are used to obtain a 3D reconstruction of an object (see section 6 for a discussion on 3D imaging and ranging). Despite being commonly treated as separate research fields, it has become obvious that, from an optical perspective, computational GI and single-pixel imaging are the same. However, it is still convenient to maintain a distinction between the two, where single-pixel imaging (or single-pixel cameras) often use a structured detection scheme, and compressed sensing, whereas computational GI often uses a structured illumination scheme. The difference between these two schemes can be demonstrated by interchanging the locations of the light source and the detector in the setups illustrated in Fig. 2 and Fig. 3.

4. Modulation schemes

As previously shown in Fig. 1, there are several choices regarding the modulation technologies used to produce the patterns for either structured detection or structured illumination single-pixel imaging systems. A useful table listing the advantages and disadvantages of various elements of single-pixel imaging systems can be found in Ref. [54].

4.1. Pseudothermal

A source of pseudothermal light can be generated by passing a laser beam through a rotating ground-glass diffuser [25]. In the case of a static diffuser a speckle pattern is generated, resulting from the diffusively transmitted light that undergoes constructive and destructive interference in different spatial regions. When the diffuser is rotated, the intensity cross-section of the resulting optical beam varies with time. In order to avoid repetition of the light field every full rotation of the diffuser, transmission through a turbid solution of microspheres can be used to further spatially randomize the pattern [48]. The pseudothermal light that emerges compares in its coherence properties to the light of an actual thermal source such as an LED [55]. As discussed in section 3 an optical beamsplitter forms two near-identical copies of the light field which can be used as the reference and object beams in a classical GI system.

The spectral properties of the pseudothermal source is determined by the properties of the materials from which it is made. Yu et al. [11] demonstrated a GI system using a pseudothermal x-ray source produced by passing a monochromatic x-ray beam through a slit array and a movable porous gold film. More recently, Zhang et al. [12] demonstrated an ultra-low radiation x-ray GI system where the pseudothermal source is generated using a polychromatic x-ray source and a sheet of rotating sandpaper. Here, the spatial structure of the illumination is similar to the speckle pattern produced using a laser and rotating ground-glass diffuser, however, this is now due to absorption rather than laser interference. The characteristics of these speckle-like features are determined by the size and transmission properties of the silicon carbide grains in the sandpaper.

4.2. Liquid crystal spatial light modulators

Pseudothermal light beams can also be generated by applying controllable random phase masks, $\phi_r(x, y)$, using a liquid crystal spatial light modulator (LC-SLM), a computer-controlled diffractive optical element which has enhanced a number of research fields in recent years. LC-SLMs impose a prescribed amount of phase shift at each pixel in an array by varying the local optical path length. Typically, this is accomplished by controlling the local orientation of the molecules in a nematic liquid crystal layer covering an array of electrodes. These are generally reflective

devices and they have an associated diffraction efficiency, fill factor and overall reflectivity, which determines their overall optical efficiency. Examples of computational GI schemes using an LC-SLM with a single-pixel detector can be found in Shapiro [26] and Katz et al. [27].

4.3. Digital micromirror devices

Digital micromirror devices (DMDs), consisting of an array of hundreds of thousands of individually addressable micromirrors, were originally developed for the display industry [56]. They offer a method of modulating light which is fast and works over a broad range of wavelengths. Micromirrors can be individually oriented at $\pm 12^\circ$, with respect to the plane of the array, by displaying a binary pattern on the DMD. The result is that light normally incident on the DMD is redirected into two paths at $\pm 24^\circ$ respectively i.e. $2 \times \pm 12^\circ$, as illustrated in Fig. 2. In a typical single-pixel camera configuration the DMD is implemented as a programmable binary transmission mask where only the path of light arising from the micromirrors in the “on” state, corresponding to a value of “+1” in the binary pattern, is transmitted and the other path, corresponding to “0”, is blocked. This can be used to structure the detected image intensities and is commonly referred to as structured detection (as previously illustrated in Fig. 2). Alternatively, the DMD can be used with a light source to project intensity patterns onto a scene, commonly referred to as structured illumination (as previously illustrated in Fig. 3).

The use of light can be optimised by measuring the light in both the positive and negative directions of the mirror tilt. Using two detectors in this manner enables a differential measurement to be performed. However, this is more commonly achieved using just one detector by displaying a pattern which is immediately followed by its contrast inverse. In addition, the background signal noise is noticeably reduced with the differential scheme, especially in the presence of illumination noise. Of course, this differential approach comes at the cost of doubling the number of binary patterns that need to be displayed on the DMD and hence, an increase in the time required to collect the data and reconstruct the image. However, DMDs are commercially available having binary pattern display rates of 22.7 kHz, which for relatively low-resolution applications allows near-video rate image reconstruction on a standard performance computer [8].

The superior modulation rate and broad wavelength response of available DMD systems, in comparison to those based on liquid crystal technology, make DMDs the common choice for use in computational imaging systems. They are particularly compatible with multi-spectral applications, where a small number of different detector types are used to measure the correlated intensities, assuming that the broad spectral response of the DMD is greater than the combined spectral responses of the individual detectors. The aluminium micromirrors of the DMD are compatible with light from the UV to the IR. However, careful consideration is required when operating at the longer wavelengths due to diffraction effects arising from the pitch of the micromirror array elements, typically 10-15 μm for many devices. Despite this limitation, standard DMDs can be used to indirectly modulate THz beams for THz single-pixel imaging (even when using wavelengths typically hundreds of μm). Stantchev et al. [57] used a DMD to spatially modulate an 800 nm pump beam which was imaged onto the back of a silicon wafer in order to modulate the THz beam. THz imaging systems have potential for applications in non-invasive imaging of concealed structures, such as in the semiconductor manufacturing industry. DMDs have also been demonstrated in other novel imaging applications. Gao et al. demonstrated compressed ultrafast photography (CUP) by using a DMD with a streak camera and based on compressed sensing [58], achieving single-shot CUP at one hundred billion frames per second.

4.4. LED arrays

The limited frame rate of many single-pixel cameras and computational GI systems has limited their use for dynamic imaging applications. Following the early demonstrations of single-pixel

imaging, many research groups have utilized compressive sampling in order to significantly reduce the number of mask patterns required to successfully reconstruct an image. However, there is still a computational cost associated with compressive sampling schemes. Recently Xu et al. [32] demonstrated a computational ghost imaging system that could continuously capture 32×32 pixel images of a dynamic scene at a rate of 1000 fps, approximately two orders of magnitude larger than other existing ghost imaging systems, by utilising an LED array for high-speed structured illumination. This was achieved by utilising the very fast ($< 1 \mu\text{s}$) switching time of the LEDs, along with the symmetry present in the Hadamard basis set that was used.

5. Pattern choice

For a camera to image without using a pixelated sensor array we need to apply a series of masks to acquire the spatial information. In the early days of television this was achieved using a physical mask, a rotating Nipkow disc consisting of a spiral arrangement of holes [59]. The signal was measured as each of the holes rotated past the scene, and line-by-line this would construct an image [34,60]. The modern version of applying a mask is to use a DMD or SLM (see section 4), enabling the mask to be dynamic and displaying a set of carefully chosen masks. The mask set, or sampling basis, can be chosen from a range of options for making a single-pixel image measurement. The simplest method would be to emulate that of early television and measure a single area per-pixel, effectively raster scanning a single pixel over the scene; this per-pixel measurement works well with high light levels but is an inefficient use of the available light [5].

GI using individual pairs of entangled photons takes many hundreds of individual measurements to form an image [42]. As discussed in section 3, it is possible to project a pseudothermal light field consisting of speckle patterns to perform classical GI but this again takes many measurements to produce a useful image [24]. Section 7 discusses a range of strategies to perform compressive sensing in order to reduce the number of measurements required when using a random basis set. In contrast, an orthogonal basis set systematically samples the scene to acquire the image such that an image is broken down into its component spatial frequencies and recorded.

5.1. Random binary

Sampling with speckle patterns could be simulated with random patterns [27,61]. These random patterns could be grey-scale values, however if fast acquisition is desired then a DMD will be able to project a series of binary patterns at much faster rates. These random patterns will reconstruct the image [26], though it can take a very large number of samples to produce a low noise image [30]. A more efficient sampling can be performed by using differential measurements, taking measurements for both sides of the DMD by using two sensors and subtracting the measured signals. The method can also be improved by using exactly half the pixels for each measurement [29]. An example of the output of these differential measurements is shown in Fig. 4.

5.2. Hadamard transform

The Hadamard matrix can be used as a basis for various sensing and imaging applications, such as recording the spatial frequencies of an image [62,63] or multiplexing the direction of illumination in a scene [64,65]. In the case of a single-pixel camera, the use of a Hadamard basis to sample the image was demonstrated by Duarte et al. [5]. The Hadamard patterns are orthogonal with binary values of +1 or -1, the Hadamard matrix is derived from the initial matrix H_2 to produce any $4k$ sized matrix.

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad (2)$$

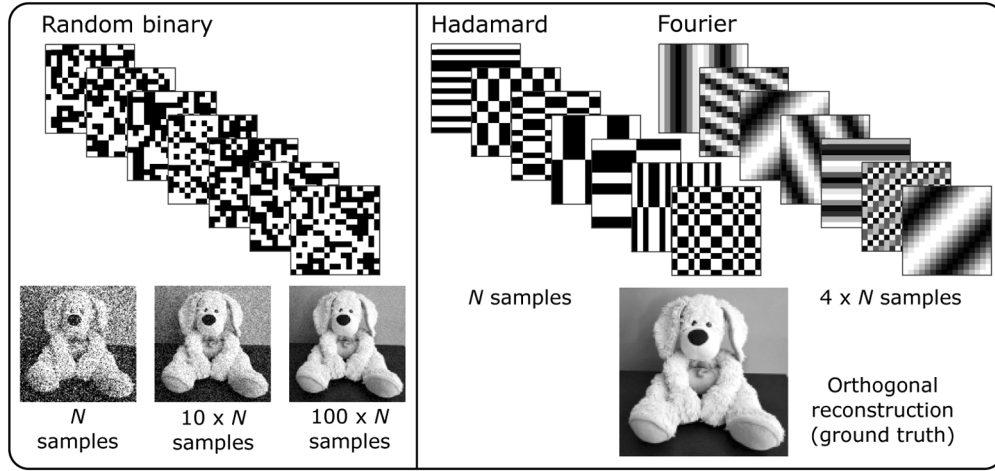


Fig. 4. A comparison of the patterns used for single-pixel imaging to reconstruct a 16×16 image. The random binary patterns will require a greater number of measurements to reconstruct an image with high accuracy. Examples are shown of an image (128×128) with the number of samples equal to the number of pixels (N), and also for 10 and 100 times the number of pixels. Patterns were measured with a differential measurement using a 50% split in the binary selection. The Hadamard patterns and Fourier patterns are orthogonal and fully sample the image, a noiseless sampling will reproduce the ground truth image with N and $4 \times N$ patterns respectively.

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix} \quad (3)$$

These matrices are their own transpose such that $HH^T = nI_n$, meaning that image reconstruction can be performed without matrix inversion. For image processing the naturally ordered Hadamard matrices can use the Walsh-Hadamard transform to calculate the result of a Hadamard matrix multiplied by a vector, with the existence of a fast Walsh-Hadamard transform (FWHT) making minimal demand on the computation required [66]. The imaging masks are created for a N pixel image (i.e. $\sqrt{N} \times \sqrt{N}$) by using the Hadamard matrix of size $N \times N$. Each row is reshaped to be the size of the image and the signal can be measured for that matrix. An example of these patterns is shown in Fig. 4. The final image is reconstructed as the Hadamard matrix multiplied by the vector of measured signals S to produce a one-dimensional vector of the output image O that requires to be reshaped into the 2D image,

$$O = H \cdot S \quad (4)$$

The orthogonality of the Hadamard basis is maintained when the elements of each of the patterns is either +1 or -1, rather than the +1 or 0 that can be displayed on the DMD. Therefore, the differential signal acquisition approach is commonly used when displaying Hadamard patterns. To sample the scene a measurement must be performed for both the +1 and -1 Hadamard values; this can be performed using either the two detector or single detector differential measurement scheme as discussed in section 4.3. This differential measurement removes any offset in the image due to background light, or slow variations in the illumination source brightness. However, this differential approach comes at the cost of requiring twice the number of patterns to be displayed on the DMD.

To demonstrate how the number of patterns can be reduced and still recreate an image we can consider what happens when we reduce the frequency range of the Hadamard patterns used. The

frequency spectrum can be determined by the number of changes in the pattern, how many times the image changes between -1 and $+1$ (for the Hadamard patterns this value is the same for all the rows and for all the columns in a single pattern). Figure 5 shows this measurement in the x and y directions, enabling a frequency spectrum to be produced with the signal measured for each pattern. The Hadamard spectrum shows the zero-frequency component in the top left and the maximum frequency in the lower right. The plot demonstrates that with orthogonal patterns the number of patterns used to capture the image can be reduced and will change the resulting image quality. The difference between the ground truth and the produced image can be measured using the mean squared error (MSE) as the difference between the ground truth intensity image I_{GT} and the reconstructed image I , defined as

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I_{GT}(i,j) - I(i,j))^2. \quad (5)$$

From this the Power signal-to-noise ratio ($PSNR$) in decibel (dB) is defined as

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_{I_{GT}}^2}{MSE} \right). \quad (6)$$

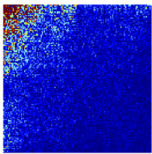

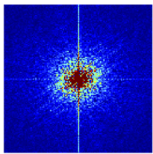

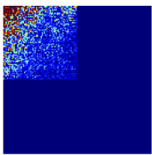

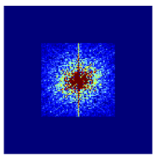

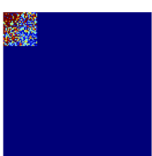

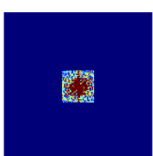



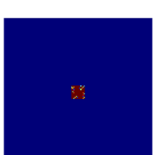

| Patterns used | Hadamard Spectrum | Reconstructed Image | Fourier Spectrum | Reconstructed Image |
|---------------|---|--|--|--|
| 100% |  |  Ground truth |  |  Ground truth |
| 25% |  |  PSNR = 59.7 dB |  |  PSNR = 65.9 dB |
| 5% |  |  PSNR = 48.5 dB |  |  PSNR = 53.0 dB |
| 1% |  |  PSNR = 40.0 dB |  |  PSNR = 42.8 dB |

Fig. 5. The sampling frequencies for the Hadamard and Fourier sampling methods [67], based on a 128×128 image. Reducing the number of patterns used to reconstruct the image produces a lower quality image, as shown by the $PSNR$ values for each image.

These calculations are performed for the different numbers of patterns used, with a comparison of the Hadamard and Fourier basis shown in Fig. 5. A square cut-off is used to reduce the number

of Hadamard patterns to reconstruct the image. The relation being that a significant reduction of the patterns can be made, which effectively reduces the number of pixels in the image.

5.3. *Fourier basis*

Other sampling schemes have been based on Fourier encoding of the pattern set [31] with further work showing some advantages over the Hadamard sampling method [67]. Whereas the Hadamard patterns use arrays of binary values, the Fourier patterns use gray-scale values. These gray-scale values can be produced with a DMD by dithering the mirrors during acquisition, or by using a high-resolution DMD and having a “super-pixel” of several mirrors, where the light gradient is controlled by the ratio of the mirrors in the “on” and “off” states. The same frequency is displayed with different phase values, with methods varying from using 3 or 4 phase values equally spaced between 0 and 2π . For a square image consisting of N pixels the patterns are created for the spatial frequencies 0 to $(\sqrt{N} - 1)$ in both the x and y dimensions, with the frequencies u and v respectively. The pattern $P(u, v)$ is generated for the image as

$$P(u, v) = \cos \left(2\pi \left(\frac{ux}{\sqrt{N}} + \frac{vy}{\sqrt{N}} \right) + \phi \right) \quad (7)$$

An example of these patterns is shown in Fig. 4. The intensity signal is insufficient to make an image reconstruction, a measurement to acquire the phase is performed by changing the phase term, ϕ , in Eq. (7). The Fourier spectrum component $\mathcal{F}(u, v)$ for the spatial frequencies u and v defined for four values of ϕ is

$$\mathcal{F}(u, v) = (D_\pi - D_0) - i(D_{3\pi/2} - D_{\pi/2}) \quad (8)$$

where D_ϕ is the intensity measurement for the signal for each of the patterns P . An inverse Fourier transform applied to the Fourier spectrum will reconstruct the image. The Fourier spectrum allows for a sensible filtering of the number of patterns needed to reconstruct the image. The effect of reducing the number of sampling frequencies used to reconstruct an image is shown in Fig. 5. It has been demonstrated that changing the shape of the cut-off, from a square to a circle or diamond, can produce different fidelity in the image reconstruction while using the same number of patterns [67].

6. 3D imaging and ranging

Three-dimensional (3D) imaging and ranging is a research field that supports a wide range of applications including object detection and classification, surface mapping and 3D situation awareness for autonomous vehicles. Within the field of computational GI, two main techniques are used, each having their advantages and drawbacks which are dependent on the specific application.

6.1. *3D computational ghost imaging*

A common technique of capturing 3D images uses stereo vision [68], which is the extraction of 3D information from the images of a scene obtained from different vantage points. However, these different images need to be aligned and have the correct geometry for the technique to be successful, and is usually computationally costly. There is a wide range of articles on using multiple 2D images to estimate depth, and a review of the algorithms used can be found in Lazaros et al. [69]. An alternative technique using photometric stereo [70] captures a sequence of images, all from the same vantage point but under different lighting conditions. Each image in the sequence is lit using a different spatially separated source of illumination. These images are much easier to align provided the sequence is captured fast enough to avoid movement of the

scene between image frames. The resulting images each differ mainly in the shading profile of the scene, from which the surface normals can be estimated.

Depth information can also be estimated from the 2D images obtained from single-pixel or computational GI. A good example of this is the 3D computational imaging system demonstrated by Sun et al. [30], which uses a photometric stereo technique and multiple single-pixel detectors rather than multiple illumination sources. Multiple detectors in different positions are used to capture multiple images of a scene illuminated using a sequence of structured patterns. Similar to conventional photometric stereo imaging, the shading in each individual image appears as if it was illuminated from a different direction. Since the spatial structure of the images is determined by a single pattern projector, the images exhibit perfect pixel registration, and comparing these images allows the 3D form of the scene to be reconstructed.

6.2. Time-of-flight imaging

A time-of-flight (ToF) measurement determines the distance to an object by illuminating it with pulsed laser light and measuring the delay of the back-scattered pulses [71,72]. The distance d can be estimated by $d = \Delta t c / 2$, where Δt is the ToF and c is the speed of light. ToF can be used in a single-pixel imaging configuration to provide information on the depth of the scene while the transverse spatial resolution is provided by the single-pixel image reconstruction, allowing a 3D representation [54,73,74]. Pulsed lasers are available with a temporal resolution in the tens of picoseconds and suitable detectors can be sensitive at the single photon level. Therefore, a ToF method is compatible with long-range, high-precision depth mapping. Figure 3 illustrates a structured illumination computational GI system that also incorporates a pulsed laser illumination source for ToF measurements (similar to that described in Ref. [33]). It is important to realise that these ToF measurement schemes are also compatible with single-pixel camera configurations, such as the one illustrated in Fig. 2 and the experiments reported by Howland et al. [20,21].

In the case of regular 2D computational imaging, one average intensity measurement is recorded for each mask or projected pattern in the sequence. In the 3D scheme, a series of intensity measurements are recorded for each pattern, each element corresponding to an intensity measurement at different depths within the scene. Hence, a series of images can be reconstructed, one at each depth, forming a 3D data cube from which both the reflectivity and depth information can be extracted.

Some of the previous demonstrations of ToF single-pixel imaging (or single-pixel LiDAR) were based on photon counting (Geiger mode) detection [20,21]. However, despite the benefits of being able to image in low light conditions, photon counting detectors have the disadvantage of having an inherent dead time (typically 10s of nanoseconds) between successive measurements, reducing the total detection efficiency. This requires measurements of the back-scattered photons from many illumination pulses in order to obtain an accurate temporal response from a 3D scene. Alternatively, a high-speed photodiode can measure the temporal response from a single illumination pulse. Sun et al. [75] demonstrated a single-pixel 3D imaging system using a high-speed photodiode for measuring the time-varying response of the back-scattered light, achieving a depth accuracy of 3 mm at a detection range of 5 m.

7. Regularisation techniques

Real images are not collections of random pixel values, rather spatially adjacent pixels tend to have similar values to each other. Within traditional image processing this allows various denoising algorithms to be applied. Within single-pixel imaging denoising is also possible but the same principles allow a form of compressed sensing where the number of masks and associated measurements can be reduced to be smaller than the number of pixels in the image. Both denoising and compressed sensing can be based on a cost function for the reconstructed image which is derived from both the data and prior information, the image is then optimised

to minimise the value of this cost function. The prior information can take several forms of varying significance. At its most basic the prior can, without loss of generality, assume that all pixel values are of positive intensity. Additionally, most natural scenes when expressed in the spatial frequency domain are sparse, i.e. many of the spatial frequencies have extremely low amplitude and can be discarded. This sparsity is the basis of many image compression techniques including the ubiquitous JPEG [37], where a discrete cosine transform is used based on a fixed low-dimensionality. To avoid the need to repeatedly calculate Fourier, or similar, transforms a similar prior is to recognise that either total variation (TV) or total curvature (TC) of the intensity distribution, $O_{(x,y)}$, of a natural image is small. These regularisation functions, R , can be written as

$$R_{\text{TV}}(O) = \sum_{i=1}^N \left(\left| \frac{dI_i(O)}{dx} \right| + \left| \frac{dI_i(O)}{dy} \right| \right) \quad (9)$$

$$R_{\text{TC}}(O) = \sum_{i=1}^N \left(\left| \frac{d^2 I_i(O)}{dx^2} \right| + \left| \frac{d^2 I_i(O)}{dy^2} \right| \right) \quad (10)$$

where as before N is the total number of pixels.

This image regularisation can be considered alongside a measure of how well the reconstructed image accounts for the measured M data values. For Gaussian noise this is characterised by the average of the square of the difference between the measured and predicted signals, χ^2/M , to create a cost function, C , for the image reconstruction.

$$C = \frac{\chi^2}{M}(O) + \lambda R(O) \quad (11)$$

The quantity λ sets the balance of the reconstruction between satisfying the data or the prior (captured in the regularisation function) and is typically set at a level such that, when optimised, we have $\chi^2/M \approx 1$, ensuring that the reconstruction is the one which most satisfies the prior while still being statistically compatible with the data. As introduced above, if an image is reconstructed using the approach embodied in Eq. (1), the algorithm can be applied to cases for both $M > N$ (number of measurements exceed number of pixels) or even $M < N$ (number of measurements less than number of pixels). When $M > N$ the regularisation process is an example of denoising, when $M < N$ it is an implementation of compressed sensing [76].

The literature around compressed sensing is extensive, not just for single-pixel cameras but more widely to high-dimensional measurement systems in general. Having established a cost function upon which to optimise the reconstruction there are further subtleties as to the statistical properties of the various regularisation functions. Assuming that the regularisation term is based on the sum of many terms, e.g. the coefficients of the image spatial frequencies, one can combine these coefficients into a single number, R , in various ways. Most obvious is to calculate the sum of the squares, l_2 , of the individual coefficients, and a minimisation of this term will tend to suppress the large coefficients. However, real images of natural scenes often have a few dominant spatial frequencies and a better goal is to promote a sparsity in the spatial frequencies. To this end a more powerful regularisation is to calculate the sum of the moduli of the coefficients, l_1 [2,4]. The details behind this sophistication is beyond the scope of this article but it is worth flagging how different statistical measures yield reconstructed or denoised images with different characteristics. There is no universal optimum measure, but rather the regularisation term should be chosen that best reflects the image type. If an l_1 regularisation is to be used then it is essential to undertake this regularisation in a basis in which the typical image can be described by the smallest number of non zero coefficients (i.e. a basis in which the typical images are sparse) - which for natural scenes is often the wavelet basis.

8. Machine learning

The machine learning approach to single-pixel imaging is a newer development than many of the other single-pixel associated techniques. The approach uses deep learning in the form of a convolutional neural network (CNN) to perform the reconstruction of an image based on fewer measurements than would be required for the orthogonal sampling methods or the traditional ghost imaging techniques. The CNN exploits the development in the speed of calculations performed by graphics processor units (GPUs) to allow higher computation rates than are available on conventional computer processors. These CNNs have so far brought about breakthroughs in processing images, object identification, language processing and medical diagnosis [77,78].

An application of CNNs developed by many groups has been to make reconstructions of an image from random patterns [79–81]. This computational GI using a CNN has produced equivalent results to a regulariser method (as discussed in section 7). Sampling with random patterns is much less efficient than with a structured imaging basis. However, deep learning has allowed the sampling basis itself to be constructed to be most efficient to sample a scene. If we are able to construct the most efficient basis then efficient imaging can be made with a minimal number of measurements. This development of a sampling basis was shown by Higham et al. [23], for a 128×128 image the number of samples was made to be 666 demonstrating a 4% compression in the sampling to reconstruct 2D images at video rates. The training of the CNN produces the sampling basis (pattern set) and also the reconstruction algorithm in the form of a trained neural network, results are shown in Fig. 6. This development of a custom basis was later used to produce 3D images of a scene [33], where the deep-learned patterns were projected onto a scene and the depth image was recovered. Similar to the scheme described in section 6.2 and illustrated in Fig. 3, a pulsed laser illuminated a DMD which was used to structure the light and (using photon-counting timing electronics) a time-of-flight measurement was made to collect the depth map of the scene.

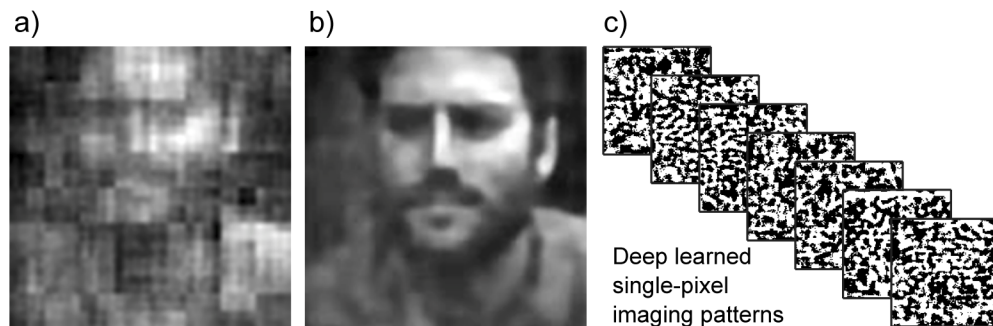


Fig. 6. An example of single-pixel imaging using deep-learning. a) The reconstruction from Hadamard sampling using a reduced number of patterns (4%). b) The reconstruction using a trained neural-network using a deep-learned patterns set. c) Examples of the deep-learned pattern sampling basis. The method applied has been presented in Higham et al. [23].

Single-pixel imaging does not necessarily need to perform a full image reconstruction to detect and classify objects. A CNN has been used to develop a low number of patterns to classify and identify very fast moving objects [82]. This technique could be extended further to enable a sensing system that feeds into a control system, such as an autonomous vehicle, meaning the creation of an image to analyse is not required for the navigation algorithms to react to a hazard presented to it, enabling much faster reaction times. Such sensing schemes are sometimes referred to as image-free classification.

Finally, using optical machine learning with a single-pixel detector may also be a possibility for image reconstruction. A diffractive neural network made up of a cascade of phase-only masks can reconstruct images without requiring the processing power of a computer [83].

9. Conclusions

We have provided a review of both single-pixel imaging and computational GI techniques and given a summary timeline of some of the main developments over the past twelve years. We have discussed some of the important aspects of the technique including, the modulation hardware, choice of pattern design, sampling strategy and the choice of detector types. While it is clear that single-pixel cameras and computational GI systems are similar in an optical sense, for the purposes of this review we found it convenient to maintain the distinction between the two. In this respect we recognise that single-pixel cameras are often based on structured detection and compressed sensing while computational GI is often based on structured illumination.

We have discussed two major advantages of single-pixel imaging which both relate to the choice of the single-pixel detector type. The first is in the design of low-cost cameras for imaging at wavelengths or multiple wavelengths outside the visible spectrum, where focal plane detector arrays are unavailable or prohibitively expensive. The second is time-resolved imaging where the time resolution of the single-pixel detector is vastly superior to that of the focal-plane array. Potential applications could be in the design of low-cost cameras for imaging at IR wavelengths, such as in gas leak detection, and for 3D imaging and ranging using LiDAR systems.

From very early on in the development of single-pixel imaging and computational GI there has been much research into ways to reduce both the data acquisition time and the image reconstruction time. Some of these techniques have been discussed in this review and include, orthogonal sampling pattern basis, compressive sensing, high-speed spatial light modulation and machine learning algorithms. Machine learning techniques have shown promise in LiDAR systems for the high-speed 3D information and ranging required for situation awareness of autonomous vehicles. It is important to realise that in such detection and classification applications it is often sufficient to detect the characteristic intensity signals without needing to reconstruct the image. Hence, fast “image-free” detection and classification is a promising research field of single-pixel imaging which could lead to an exciting new range of unique sensing technologies.

Funding

QuantIC (EP/M01326X/1).

Disclosures

The authors declare no conflicts of interest.

References

1. E. J. Candès, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Commun. Pure Appl. Math.* **59**(8), 1207–1223 (2006).
2. D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
3. M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, “An architecture for compressive imaging,” in *2006 International Conference on Image Processing*, (IEEE, 2006), pp. 1273–1276.
4. E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” *IEEE Trans. Inf. Theory* **52**(12), 5406–5425 (2006).
5. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, “Single-pixel imaging via compressive sampling,” *IEEE Signal Process. Mag.* **25**(2), 83–91 (2008).
6. P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, “Dual Photography,” *ACM Trans. Graph.* **24**(3), 745–755 (2005).
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**(5), 285–289 (2014).

8. 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**(1), 10669 (2015).
9. 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* **25**(4), 2998–3005 (2017).
10. J. Greenberg, K. Krishnamurthy, and D. Brady, "Compressive single-pixel snapshot x-ray diffraction imaging," *Opt. Lett.* **39**(1), 111–114 (2014).
11. H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, "Fourier-transform ghost imaging with hard X rays," *Phys. Rev. Lett.* **117**(11), 113901 (2016).
12. A.-X. Zhang, Y.-H. He, L.-A. Wu, L.-M. Chen, and B.-B. Wang, "Tabletop x-ray ghost imaging with ultra-low radiation," *Optica* **5**(4), 374–377 (2018).
13. W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Appl. Phys. Lett.* **93**(12), 121105 (2008).
14. D. Shrekenhamer, C. M. Watts, and W. J. Padilla, "Terahertz single pixel imaging with an optically controlled dynamic spatial light modulator," *Opt. Express* **21**(10), 12507–12518 (2013).
15. S. M. Hornett, R. I. Stantchev, M. Z. Vardaki, C. Beckerleg, and E. Hendry, "Subwavelength terahertz imaging of graphene photoconductivity," *Nano Lett.* **16**(11), 7019–7024 (2016).
16. R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *2007 IEEE radar conference*, (IEEE, 2007), pp. 128–133.
17. 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.* **4**(1), 5834 (2015).
18. V. Studer, J. Bobin, M. Chahid, H. S. Mousavi, E. Candes, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," *Proc. Natl. Acad. Sci.* **109**(26), E1679–E1687 (2012).
19. E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.* **23**(3), 969–985 (2007).
20. G. A. Howland, P. B. Dixon, and J. C. Howell, "Photon-counting compressive sensing laser radar for 3D imaging," *Appl. Opt.* **50**(31), 5917–5920 (2011).
21. G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, "Photon counting compressive depth mapping," *Opt. Express* **21**(20), 23822–23837 (2013).
22. R. I. Stantchev, D. B. Phillips, P. Hobson, S. M. Hornett, M. J. Padgett, and E. Hendry, "Compressed sensing with near-field THz radiation," *Optica* **4**(8), 989–992 (2017).
23. C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Sci. Rep.* **8**(1), 2369 (2018).
24. A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato, "Ghost imaging with thermal light: comparing entanglement and classical correlation," *Phys. Rev. Lett.* **93**(9), 093602 (2004).
25. A. Valencia, G. Scarcelli, M. D'Angelo, and Y. Shih, "Two-photon imaging with thermal light," *Phys. Rev. Lett.* **94**(6), 063601 (2005).
26. J. H. Shapiro, "Computational ghost imaging," *Phys. Rev. A* **78**(6), 061802 (2008).
27. O. Katz, Y. Bromberg, and Y. Silberberg, "Compressive ghost imaging," *Appl. Phys. Lett.* **95**(13), 131110 (2009).
28. Y. Bromberg, O. Katz, and Y. Silberberg, "Ghost imaging with a single detector," *Phys. Rev. A* **79**(5), 053840 (2009).
29. F. Ferri, D. Magatti, L. Lugiato, and A. Gatti, "Differential ghost imaging," *Phys. Rev. Lett.* **104**(25), 253603 (2010).
30. B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, "3D Computational imaging with single-pixel detectors," *Science* **340**(6134), 844–847 (2013).
31. Z. Zhang, X. Ma, and J. Zhong, "Single-pixel imaging by means of Fourier spectrum acquisition," *Nat. Commun.* **6**(1), 6225 (2015).
32. Z.-H. Xu, W. Chen, J. Penuelas, M. Padgett, and M.-J. Sun, "1000 fps computational ghost imaging using LED-based structured illumination," *Opt. Express* **26**(3), 2427–2434 (2018).
33. N. Radwell, S. D. Johnson, M. P. Edgar, C. F. Higham, R. Murray-Smith, and M. J. Padgett, "Deep learning optimized single-pixel lidar," *Appl. Phys. Lett.* **115**(23), 231101 (2019).
34. J. L. Baird, "Apparatus for transmitting views or images to a distance," (1929). US Patent 1,699,270.
35. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 21–30 (2008).
36. J. Romberg, "Imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 14–20 (2008).
37. R. Aravind, G. L. Cash, and J. P. Worth, "On implementing the JPEG still-picture compression algorithm," in *Visual Communications and Image Processing IV*, vol. 1199 (1989), pp. 799–808.
38. G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM* **34**(4), 30–44 (1991).
39. A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Process. Mag.* **18**(5), 36–58 (2001).
40. M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.* **4**(2), 445–460 (2010).
41. M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Computational Imaging V*, vol. 6498 (International Society for Optics and Photonics, 2007), p. 64980H.

42. T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, "Optical imaging by means of two-photon quantum entanglement," *Phys. Rev. A* **52**(5), R3429–R3432 (1995).
43. J. H. Shapiro and R. W. Boyd, "The physics of ghost imaging," *Quantum Inf. Process.* **11**(4), 949–993 (2012).
44. R. S. Aspden, N. R. Gemmell, P. A. Morris, D. S. Tasca, L. Mertens, M. G. Tanner, R. A. Kirkwood, A. Ruggeri, A. Tosi, R. W. Boyd, G. S. Buller, R. H. Hadfield, and M. J. Padgett, "Photon-sparse microscopy: visible light imaging using infrared illumination," *Optica* **2**(12), 1049–1052 (2015).
45. R. I. Khakimov, B. M. Henson, D. K. Shin, S. S. Hodgman, R. G. Dall, K. G. H. Baldwin, and A. G. Truscott, "Ghost imaging with atoms," *Nature* **540**(7631), 100–103 (2016).
46. A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato, "Correlated imaging, quantum and classical," *Phys. Rev. A* **70**(1), 013802 (2004).
47. R. S. Bennink, S. J. Bentley, and R. W. Boyd, "'Two-photon' coincidence imaging with a classical source," *Phys. Rev. Lett.* **89**(11), 113601 (2002).
48. F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla, and L. A. Lugiato, "High-resolution ghost image and ghost diffraction experiments with thermal light," *Phys. Rev. Lett.* **94**(18), 183602 (2005).
49. A. Gatti, M. Bache, D. Magatti, E. Brambilla, F. Ferri, and L. A. Lugiato, "Coherent imaging with pseudo-thermal incoherent light," *J. Mod. Opt.* **53**(5-6), 739–760 (2006).
50. B. I. Erkmen and J. H. Shapiro, "Ghost imaging: from quantum to classical to computational," *Adv. Opt. Photonics* **2**(4), 405–450 (2010).
51. P. Sen, "On the relationship between dual photography and classical ghost imaging," arXiv:1309.3007 (2013).
52. H. Von Helmholtz, *Handbuch der Physiologischen Optik*, vol. 9 (Voss, 1867).
53. L. Rayleigh, "XXVIII. On the law of reciprocity in diffuse reflexion," *The London, Edinburgh, Dublin Philos. Mag. J. Sci.* **49**(298), 324–325 (1900).
54. M.-J. Sun and J.-M. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: a brief review," *Sensors* **19**(3), 732 (2019).
55. K. Pieper, A. Bergmann, R. Dengler, and C. Rockstuhl, "Using a pseudo-thermal light source to teach spatial coherence," *Eur. J. Phys.* **39**(4), 045303 (2018).
56. J. B. Sampsell, "Digital micromirror device and its application to projection displays," *J. Vac. Sci. Technol., B: Microelectron. Process. Phenom.* **12**(6), 3242–3246 (1994).
57. R. I. Stantchev, B. Sun, S. M. Horne, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," *Sci. Adv.* **2**(6), e1600190 (2016).
58. L. Gao, J. Liang, C. Li, and L. V. Wang, "Single-shot compressed ultrafast photography at one hundred billion frames per second," *Nature* **516**(7529), 74–77 (2014).
59. E. Goldberg, "Nipkow disk for television," (1934). US Patent 1,973,203.
60. G. C. B. Rowe, "Television comes to the home," *Radio News* pp. 1098–1156 (1928).
61. B. Sun, S. S. Welsh, M. P. Edgar, J. H. Shapiro, and M. J. Padgett, "Normalized ghost imaging," *Opt. Express* **20**(15), 16892–16901 (2012).
62. W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard transform image coding," *Proc. IEEE* **57**(1), 58–68 (1969).
63. N. J. A. Sloane and M. Harwit, "Masks for Hadamard transform optics, and weighing designs," *Appl. Opt.* **15**(1), 107–114 (1976).
64. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "A theory of multiplexed illumination," in *Proceedings of the Ninth IEEE International Conference on Computer Vision*, vol. 2 (IEEE, 2003), pp. 808–815.
65. Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(8), 1339–1354 (2007).
66. Y. A. Geadah and M. J. G. Corinthios, "Natural, dyadic, and sequency order algorithms and processors for the Walsh-Hadamard transform," *IEEE Trans. Comput.* **C-26**(5), 435–442 (1977).
67. Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Hadamard single-pixel imaging versus Fourier single-pixel imaging," *Opt. Express* **25**(16), 19619–19639 (2017).
68. A. Boyde, "Stereoscopic images in confocal (tandem scanning) microscopy," *Science* **230**(4731), 1270–1272 (1985).
69. N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of stereo vision algorithms: from software to hardware," *Int. J. Optomechatronics* **2**(4), 435–462 (2008).
70. R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.* **19**(1), 191139 (1980).
71. T. J. Kane, C. E. Byvik, W. J. Kozlovsky, and R. L. Byer, "Coherent laser radar at 1.06 μm using Nd:YAG lasers," *Opt. Lett.* **12**(4), 239–241 (1987).
72. M.-C. Amann, T. M. Bosch, M. Lescure, R. A. Myllylae, and M. Rioux, "Laser ranging: a critical review of unusual techniques for distance measurement," *Opt. Eng.* **40**(1), 10–19 (2001).
73. A. Kirmani, A. Colaço, F. N. C. Wong, and V. K. Goyal, "Exploiting sparsity in time-of-flight range acquisition using a single time-resolved sensor," *Opt. Express* **19**(22), 21485–21507 (2011).
74. A. Colaço, A. Kirmani, G. A. Howland, J. C. Howell, and V. K. Goyal, "Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2012), pp. 96–102.
75. 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. Commun.* **7**(1), 12010 (2016).

76. C. A. Metzler, A. Maleki, and R. G. Baraniuk, "From denoising to compressed sensing," *IEEE Trans. Inf. Theory* **62**(9), 5117–5144 (2016).
77. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
78. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), pp. 4681–4690.
79. M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," *Sci. Rep.* **7**(1), 17865 (2017).
80. T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, and T. Ito, "Computational ghost imaging using deep learning," *Opt. Commun.* **413**, 147–151 (2018).
81. Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, "Ghost imaging based on deep learning," *Sci. Rep.* **8**(1), 6469 (2018).
82. Z. Zhang, X. Li, S. Zheng, M. Yao, G. Zheng, and J. Zhong, "Image-free classification of fast-moving objects using "learned" structured illumination and single-pixel detection," *Opt. Express* **28**(9), 13269–13278 (2020).
83. S. Jiao, J. Feng, Y. Gao, T. Lei, Z. Xie, and X. Yuan, "Optical machine learning with incoherent light and a single-pixel detector," *Opt. Lett.* **44**(21), 5186–5189 (2019).