

# Single-pixel imaging: An overview of different methods to be used for 3D space reconstruction in harsh environments

Cite as: Rev. Sci. Instrum. **92**, 111501 (2021); <https://doi.org/10.1063/5.0050358>

Submitted: 14 March 2021 • Accepted: 12 October 2021 • Published Online: 05 November 2021

 Carlos A. Osorio Quero,  Daniel Durini,  Jose Rangel-Magdaleno, et al.





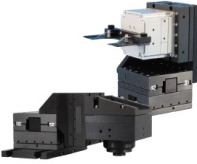
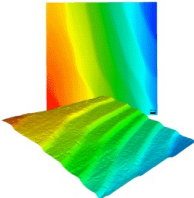
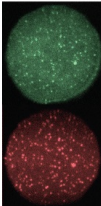
View Online



Export Citation



CrossMark

	<p>Nanopositioning Systems</p> 	<p>Modular Motion Control</p> 	<p>AFM and NSOM Instruments</p> 	<p>Single Molecule Microscopes</p> 
---	--	--	---	--

# Single-pixel imaging: An overview of different methods to be used for 3D space reconstruction in harsh environments

Cite as: Rev. Sci. Instrum. 92, 111501 (2021); doi: 10.1063/5.0050358

Submitted: 14 March 2021 • Accepted: 12 October 2021 •

Published Online: 5 November 2021



View Online



Export Citation



CrossMark

Carlos A. Osorio Quero,<sup>1</sup>  Daniel Durini,<sup>1</sup>  Jose Rangel-Magdaleno,<sup>1,a)</sup>  and Jose Martinez-Carranza<sup>2</sup> 

## AFFILIATIONS

<sup>1</sup>Digital Systems Group, Electronics Department, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), 72840 Puebla, Mexico

<sup>2</sup>Computer Science Department, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), 72840 Puebla, Mexico

**Note:** Paper published as part of the Special Topic on Advances in Measurements and Instrumentation Leveraging Embedded Systems.

<sup>a)</sup>Author to whom correspondence should be addressed: [jrangel@inaoe.mx](mailto:jrangel@inaoe.mx)

## ABSTRACT

Different imaging solutions have been proposed over the last few decades, aimed at three-dimensional (3D) space reconstruction and obstacle detection, either based on stereo-vision principles using active pixel sensors operating in the visible part of the spectra or based on active Near Infra-Red (NIR) illumination applying the time-of-flight principle, to mention just a few. If extremely low quantum efficiencies for NIR active illumination yielded by silicon-based detector solutions are considered together with the huge photon noise levels produced by the background illumination accompanied by Rayleigh scattering effects taking place in outdoor applications, the operating limitations of these systems under harsh weather conditions, especially if relatively low-power active illumination is used, are evident. If longer wavelengths for active illumination are applied to overcome these issues, indium gallium arsenide (InGaAs)-based photodetectors become the technology of choice, and for low-cost solutions, using a single InGaAs photodetector or an InGaAs line-sensor becomes a promising choice. In this case, the principles of Single-Pixel Imaging (SPI) and compressive sensing acquire a paramount importance. Thus, in this paper, we review and compare the different SPI developments reported. We cover a variety of SPI system architectures, modulation methods, pattern generation and reconstruction algorithms, embedded system approaches, and 2D/3D image reconstruction methods. In addition, we introduce a Near Infra-Red Single-Pixel Imaging (NIR-SPI) sensor aimed at detecting static and dynamic objects under outdoor conditions for unmanned aerial vehicle applications.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0050358>

## I. INTRODUCTION

This work aims at presenting a concise review of the different applications reported in the literature based on the single-pixel imaging principle, which among other advantages could also enable the use of single InGaAs-based photodetectors (or InGaAs line-sensors) in low-cost Near Infra-Red Single-Pixel Imaging (NIR-SPI) systems as the one proposed in this work. The latter is an alternative sensing paradigm for 2D/3D vision applications, which are normally based on stereo-vision principles using active pixel (RGB or RGB-D) sensors operating in the visible (VIS) part of the spectra

(in the wavelength range between 450 and 750 nm) or apply the time-of-flight (TOF) principle, as is the case in, e.g., Laser Imaging Detection and Ranging (LIDAR) systems, which use active laser-based pulsed illumination or LED-based continuous-wave illumination in the near infra-red (NIR) part of the spectra with 850 nm or 905 nm wavelengths, to mention just a few. Silicon yields an extremely low quantum efficiency (QE) in the NIR part of the spectra (of below 10%) and completely vanishes beyond the 1107 nm wavelength,<sup>1</sup> which negatively affects its detection performance of the NIR active illumination used. Unfortunately, silicon-based imagers simultaneously excel at detecting background illumination (mostly

in the visible part of the spectra), which produces huge amounts of photon-shot-noise and gravely diminishes the systems' ability to detect emitted NIR laser pulses (or LED originated radiation) that first need to reach the different objects in the scene in front of the imaging system, then get reflected by them, and are finally detected by using the image sensor placed alongside the active illumination source. If micrometer-sized particles are present in the environment, Rayleigh and other scattering effects take place that prevent the emitted and reflected pulses from reaching the photosensor. Therefore, in outdoor applications, the operating limitations of silicon sensor-based systems under harsh weather conditions, especially if relatively low-power active illumination is used (normally limited by eye-safety regulations in the case of pulsed laser sources), are quite obvious. If longer wavelengths (e.g., of 1550 nm) for the active illumination in these systems are applied, InGaAs-based photodetectors, considering their high QE (>80%) in the wavelength range between 900 and 1600 nm, become the technology of choice. If low-cost solutions based on this approach are pursued, then using a single InGaAs photodetector or an InGaAs line-sensor is the only affordable choice, and the principles of Single-Pixel Imaging (SPI) and Compressive Sensing (CS) acquire a paramount importance. The latter has motivated us to present a review of the principles and main techniques behind SPI systems, their evolution over time, and their potential applications. We also introduce a new NIR-SPI device oriented toward drone applications.

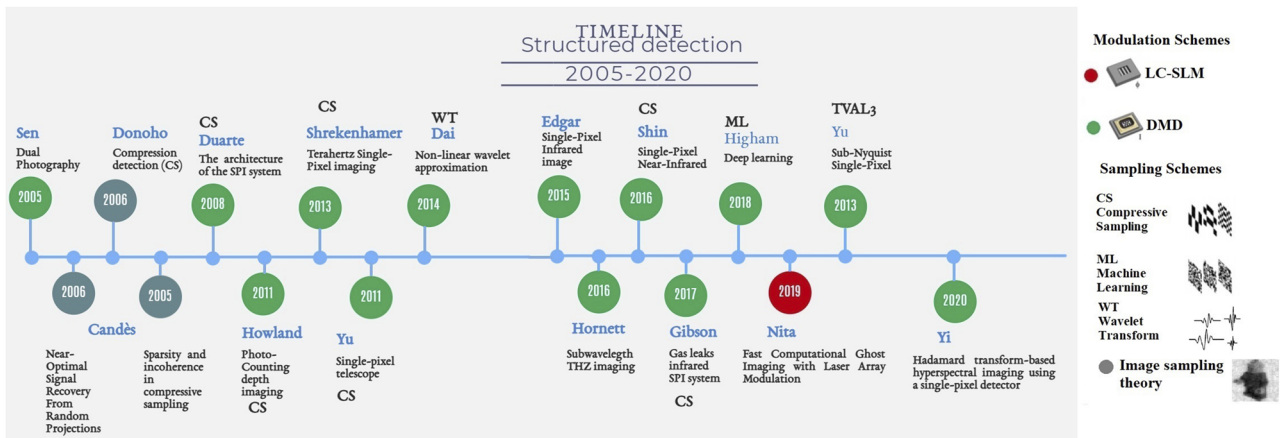
At first glance, having a single pixel may seem counterproductive as today's cameras incorporate several million pixels at quite acceptable costs. However, light sensors with a pixelated structure are not always necessary, and the technology of the only pixel has some advantages that stand out compared to cameras based on a series of sensors (architecture of CCD or CMOS cameras<sup>1</sup>). First of all, single-point detectors generally have high efficiency and can detect weak changes in light intensity. Therefore, a single-pixel image (SPI) has become a modern paradigm of image acquisition through data compression. A single-pixel camera (SPC) architecture consists of just two elements: a spatial light modulator and a single-point detector. The key idea here is to modulate the observed image with a specific pattern and collect the single detector's corresponding measurement. The desired image can be restored by post-processing a sequence of measurements made with various patterns.

The first SPI concept was originated from the idea of modulating a light field and collecting this modulated light by using a single photodetector. This concept was documented in 1982 by Ben-Yosef and Sirat,<sup>2</sup> a work in which the authors proposed using the elastic piezoelectric optical effect taking place in crystals for light modulation, forcing the light output to remain proportional to the Fourier transform of the photographed object. However, back then, the construction of small and numerous crystals was not as accessible as today, and the authors only presented a proof of concept using a few crystals to restore the image of an object. Almost 25 years later, at Rice University, the first SPC was proposed and its performance was successfully demonstrated based on the pioneering idea of compression detection (CS) proposed by Donoho<sup>3</sup> in 2006; in the same year, it was also proposed in parallel by Takhar *et al.*<sup>4</sup> Their approach was based on the use of random patterns to reconstruct an image using a minimization algorithm applied for the SPC camera. By 2008, the architecture of the SPI system was redefined as reported in the work of Duarte *et al.*,<sup>5</sup> consisting, in this case,

of a light source element, a spatial light modulator (SLM), and a detector element [single-pixel detector (SPD)]. They presented their image recovery method using the compressed sensing approach proposed by Donoho<sup>3</sup> two years before. Duarte proposed the first color image processing method in 2008, and for this, an RGB filter was used in conjunction with a single photodiode, and the measuring method consisted of performing three consecutive measurements using a different RGB filter each time to form the color pattern. This approach was improved in 2013 by Welsh *et al.*<sup>6</sup> by introducing a dichroic beam splitter, which allowed for decomposing the white light based illumination into three separate outputs (red, green, and blue) and proposed placing one different photodiode at each of the three outputs to collect the resulting light. This process allowed for restoring an image by using each separate color channel output, respectively, and obtaining a unified color image by combining these outputs. Other types of applications using the single-pixel imaging approach include the creation of 3D images,<sup>7</sup> video streams, hyper-spectral applications including a combination of infrared images and images obtained in the visible range, or including 3D images obtained by applying the TOF principle or even radar systems. A timeline showing all these developments can be observed in Fig. 1.

Following the developments mentioned, new methods appeared that were aiming at more efficient and accurate processing of photodetector outputs used for image reconstruction, alongside new strategies used to generate illumination patterns and capture them more efficiently. Sun *et al.* proposed in 2013 one of the first approaches that enabled the generation of 3D images based on the SPI principle.<sup>7</sup> They proposed using a projector to illuminate the scene with random patterns and place four photodetectors at different angles that would detect the light reflected from the objects in the scene and simultaneously measure the time it took for each generated illumination pattern to reach the scene, be reflected by the objects in it, and finally reach and be detected by them. This method was based on the TOF principle and the approach proposed by Howland *et al.* in 2013,<sup>11</sup> applied in an SPC. In 2014, Dai *et al.*<sup>15</sup> proposed an adaptive basic scanning strategy based on the illumination patterns generated by a method using the wavelet transform and an image recovery approach using the wavelet inverse transform. The latter soon became a reference that will serve as a strategy for the generation of depth maps for systems that work in combination with the time-of-flight principle. That development reported by Dai *et al.* inspired Zhang *et al.* in 2015<sup>30</sup> to use illumination patterns generated by using the Fourier transform and perform image recovery by applying the inverse Fourier transform to the photodetector output signals, a concept that will set the basis for the subsequent video sequence generation and 3D applications using the SPI principle.<sup>33</sup> These applications take advantage of temporal redundancy to reduce the number of processing steps to reconstruct depth maps and so reduce the processing time required for each individual image, enabling video streaming.

The latest advances in SPI-type systems focused on approaches that could avoid the use of lenses for light structuring, as used in PicoCam<sup>34</sup> and FlatCam<sup>35</sup> developments, presented in 2017; the implementation of emerging carbon nanotube<sup>36</sup> and graphene technologies, respectively, used in photodetectors, as proposed in 2017 and 2018;<sup>7</sup> LED array illumination substituting the much more



**FIG. 1.** Timeline showing different developments based on the single-pixel imaging approach using structured detection, referred to as computational ghost imaging, showing the different modulation technologies developed and employing the sampling scheme based on compressed sensing (CS), machine learning (ML), and wavelet transform (WT) approaches, respectively, between 2005 and 2020. The following references are shown: Sen *et al.*,<sup>8</sup> Candes and co-workers,<sup>9,10</sup> Donoho,<sup>3</sup> Takhar *et al.*,<sup>4</sup> Duarte *et al.*,<sup>5</sup> Howland *et al.*,<sup>11,12</sup> Shrekenhamer *et al.*,<sup>13</sup> Yu *et al.*,<sup>14</sup> Dai *et al.*,<sup>15</sup> Radwell *et al.*,<sup>16</sup> Hornett *et al.*,<sup>17</sup> Gibson *et al.*,<sup>18</sup> Higham *et al.*,<sup>19</sup> Nitta *et al.*,<sup>20</sup> Yu,<sup>21</sup> and Yi *et al.*<sup>22</sup>

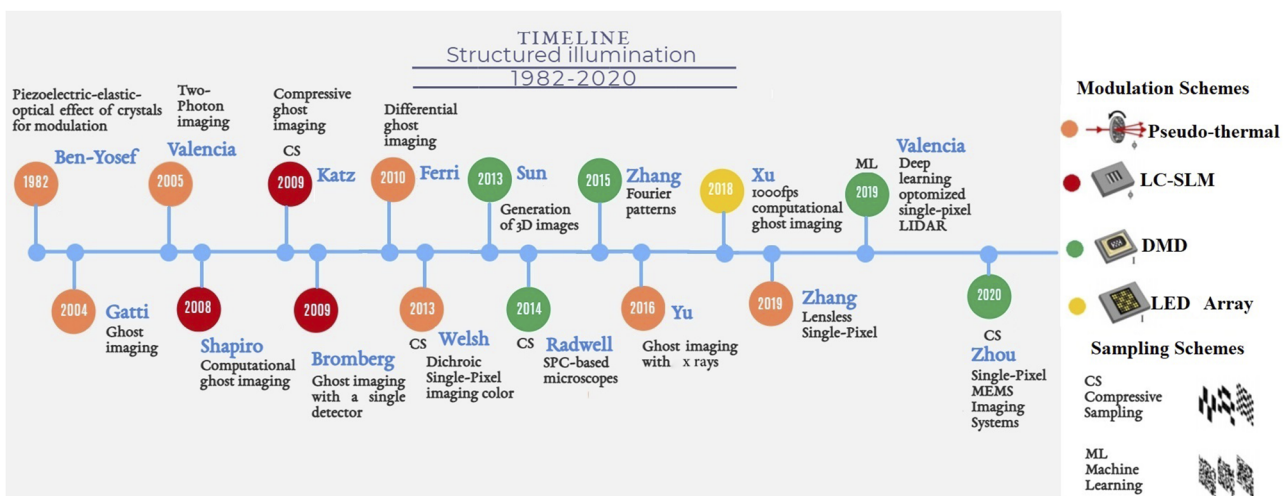
expensive SLMs proposed in 2018;<sup>37</sup> deep learning approach for image recovery proposed in 2019;<sup>29</sup> or hyperspectral imaging using the single-pixel approach proposed in 2020.<sup>22</sup> Figures 1 and 2 show a timeline with different developments mentioned, used in single-pixel imaging systems, including modulation technologies, sampling, and processing schemes, respectively, developed between 1982 and 2020.

## II. BASIC PRINCIPLE OF SINGLE-PIXEL IMAGING

Single-pixel imaging is based on the principle of spatial modulation of light, which relies on the projection of a sequence of structured illumination patterns, generated using light modulation

devices, such as SLM, Digital Micromirror Devices (DMDs), or other similar modulators (see Fig. 1), onto the object to be imaged and the detection of this modulated light reflected by the object, using a lens system for focusing, by using a single photodiode that delivers an output voltage signal equivalent to the amount of light detected. The relation between the structured illumination pattern and the light signal reflected from the imaged object and finally detected by using the photodiode in each measurement can be expressed through the following equation:<sup>38</sup>

$$S_i = \alpha \sum_{x=1}^M \sum_{y=1}^N O(x, y) \Phi(x, y). \quad (1)$$



**FIG. 2.** Timeline showing the different developments based on single-pixel imaging with structured illumination, referred to as computational ghost imaging, showing different modulation technologies proposed and different sampling approaches used based on compressed sensing (CS) or machine learning (ML) between 1982 and 2020. The following references are shown: Ben-Yosef and Sirat,<sup>2</sup> Gatti *et al.*,<sup>23</sup> Valencia *et al.*,<sup>24</sup> Shapiro,<sup>25</sup> Katz *et al.*,<sup>26</sup> Bromberg *et al.*,<sup>27</sup> Ferri *et al.*,<sup>28</sup> Welsh *et al.*,<sup>6</sup> Sun *et al.*,<sup>7</sup> Radwell *et al.*,<sup>29</sup> Zhang *et al.*,<sup>30</sup> Yang *et al.*,<sup>31</sup> and Xu *et al.*<sup>32</sup>

Here,  $(x, y)$  are the spatial coordinates,  $O$  denotes the reflectivity of the illuminated object,  $\Phi_i$  stands for the  $i$ -th structured pattern emitted in the sequence used,  $S_i$  is the  $i$ -th single-pixel (photodetector) measurement output corresponding to  $\Phi_i$ , and  $\alpha$  is a factor describing the optoelectronic response of the photodetector (see Fig. 3). The final reconstructed image size, which applies to both, the object being imaged and the amount of patterns projected, is defined by  $MN$  pixels. From the definition of each of the structured illumination patterns projected and the output voltage signal delivered by the photodiode in response to each of those particular patterns being reflected from the illuminated object, a final virtual image of the illuminated object can be recovered. Considering that the reconstructed image  $I$  is proportional to the object reflectivity  $O$ , the reconstructed object image can be recovered through the application of the following equation:<sup>38</sup>

$$I(x, y) = \alpha \sum_{x=1}^M \sum_{y=1}^N S_i \Phi_i(x, y). \quad (2)$$

Here, the reconstructed image is an expression as the inner product of the output voltage signal obtained in each measurement and the structured pattern applied to obtain it. The spatial light modulation in single-pixel imaging can be carried out following two different schemes:<sup>39</sup> (1) a structured illumination scheme named front modulation [depicted in Fig. 4(a)] and (2) a structured detection scheme named back modulation [depicted in Fig. 4(b)].

In the structured illumination scheme (see Table I), the light modulation device is placed between in front of the object to be imaged. Here, the light source (normally white) illuminates the SLM that generates a particular illumination pattern that is then projected onto the target object through a lens. In this way, structured illumination stimuli are generated, and the back-scattered light, reflected by the illuminated object created by it, is finally detected by using the single-pixel detector (SPD). In the structured illumination detection scheme, the light modulation device is placed behind the target

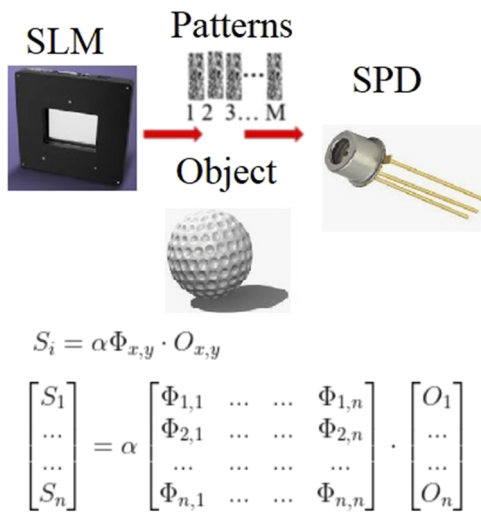


FIG. 3. The single-pixel imaging setup: SLM, patterns, and the measurement signal of SPD.

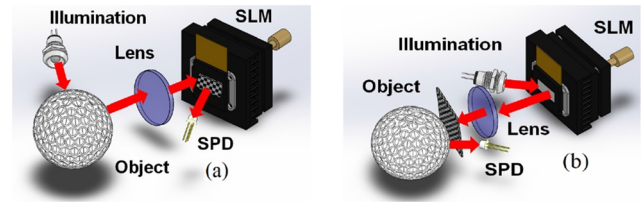


FIG. 4. The two different approaches applied to SPI: (a) Front modulation: the object is illuminated by using a light source and the light reflected by it gets directed by means of a lens onto an SLM device where a sequence of patterns is loaded. Each specific pattern enables certain pixel-mirrors to reflect light and others not to do it. The light reflected from the SLM presenting a specific pattern is finally detected by using a photodetector [or single-pixel detector (SPD)], and its output voltage signals generated by each particular illumination pattern are processed later to recover the image. (b) Back modulation: the SLM device programmed with the specific sequence of patterns is illuminated by using the light source, and the light reflected by each individual pattern is then projected by means of a lens onto the object. The object generated reflected image is finally captured by using the SPD, and its output voltage signals generated by each particular illumination pattern are processed later to recover the image.<sup>39</sup>

object. The characteristic of the detection components (light modulation device and the SPD) determines the field of view of the reconstructed image. The spatial resolution and the overall quality of each reconstructed single-pixel image obtained in this manner depend on the following three criteria:

1. The type of structured patterns used in the spatial light modulation (see Table III).
2. The sampling and scanning strategy used.
3. The image reconstruction algorithm (see Tables II and V).
4. Reconstruction quality vs running time (see Fig. 5).

There are seven scanning and sampling strategies documented in the literature that have been proposed to be used for reconstruction of single-pixel images: (1) Computational Ghost Imaging (CGI), (2) Compressive Sensing Ghost Imaging (CSGI), (3) Hadamard Single-Pixel Imaging (HSI), (4) Fourier Single-Pixel Imaging (FSI), (5) Binary Fourier Single-Pixel Imaging (BFSI), (6) Wavelet (WL) single-pixel imaging, and (7) Machine Learning (ML) single-pixel imaging.

TABLE I. Summary of single-pixel imaging system architectures.<sup>41</sup>

Architecture	Advantage	Disadvantage
Focal plane modulation	Active or passive imaging	Limited choice on modulation
Structured light illumination	More choices for active illumination	Active imaging only
Rotating ground glass	High power endurance and cheap	Not programmable and random modulation only
Customized diffuser	High power endurance and can be customized	Not programmable and complicated manufacturing

TABLE II. Summary of the reconstruction algorithm.<sup>41</sup>

Algorithm	Advantage	Disadvantage
Orthogonal sub-sampling (TVAL3 <sup>42</sup> and NESTA <sup>43,44</sup> )	Not computationally demanding	Requires a specific prior
Compressed sensing (CS) (orthogonal matching <sup>45</sup> )	A computational overhead	Need only a general sparse assumption
No iteration matrix inversion conventional correlation <sup>27</sup> and DGI <sup>28</sup>	Required a low running time	Required high measurement numbers

### A. Computational ghost imaging (CGI)

Computational Ghost Imaging (CGI) is a technique that acquires spatial information about an unknown target by generating random patterns for spatial light modulation<sup>28</sup> [depicted in Fig. 6(a)]. Typically, the random patterns are binary so that CGI can take advantage of high-speed binary pattern generation given by a digital micro-mirror device (DMD). The CGI uses a correlation-based algorithm for image reconstruction expressed through Eq. (3),<sup>46</sup> so it needs many measurements (that is,  $n = N \times M$ ) to produce a good-quality image. The CGI, due to the simplicity of deployment of low-cost, robustness against noise and scattering, its ability to operate over a long spectral range, and its capacity of inherent encryption of patterns,<sup>23</sup> is widely used in many SPI applications.<sup>24,28</sup> The images reconstructed by applying this approach can be obtained by applying the following equation (see Algorithm 1):

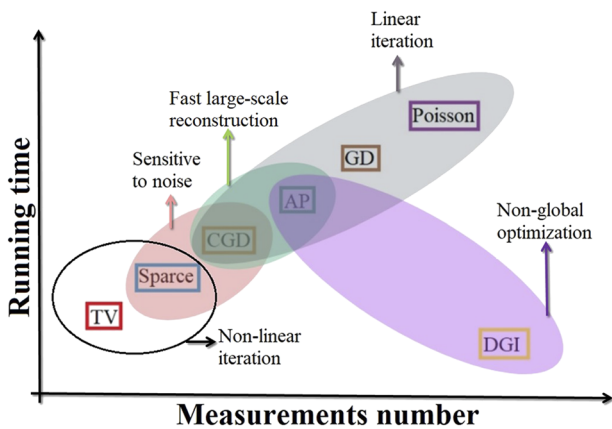


FIG. 5. Comparison diagram of different SPI algorithms that show the performance based on the measurement number to make a reconstruction quality vs running time. The red dotted circle represents the sparse representation and CGD algorithms sensitive to measurement noise, the purple dotted circle represents non-global optimization AP and DGI algorithms, and the green dotted circle represents AP and CGD algorithms that take the least running time for large-scale reconstruction.<sup>40</sup>

$$I = \frac{1}{n} \sum_{i=1}^n \left( S_n - \frac{1}{n} \sum_{i=1}^n S_i \right) \left( \Phi_n - \frac{1}{n} \sum_{i=1}^n \Phi_i \right). \quad (3)$$

In Eq. (3),  $n$  is the number of structured patterns and  $I$  stands for reconstructed image equivalent to the weighted sum of the structured patterns.

### B. Compressive sensing ghost imaging (CSGI)

CSGI is an improvement of the technique of CGI based on Compressive Sensing (CS),<sup>3</sup> which allows us to recover an image with fewer single-pixel measurements ( $n \gg MN$ ). It was proposed to improve the low data efficiency problem present in CGI. CSGI attempts to recover the object image  $I$  with a size of  $MN$  by solving a set of  $n$  equations using different sparsities of an image to solve  $G$  in the following equation:<sup>22</sup>

$$G(I) = \sum_{x=2}^M \sum_{y=2}^N \{ [I(x, y) - I(x-1, y)]^2 + [I(x, y) - I(x, y-1)]^2 \}. \quad (4)$$

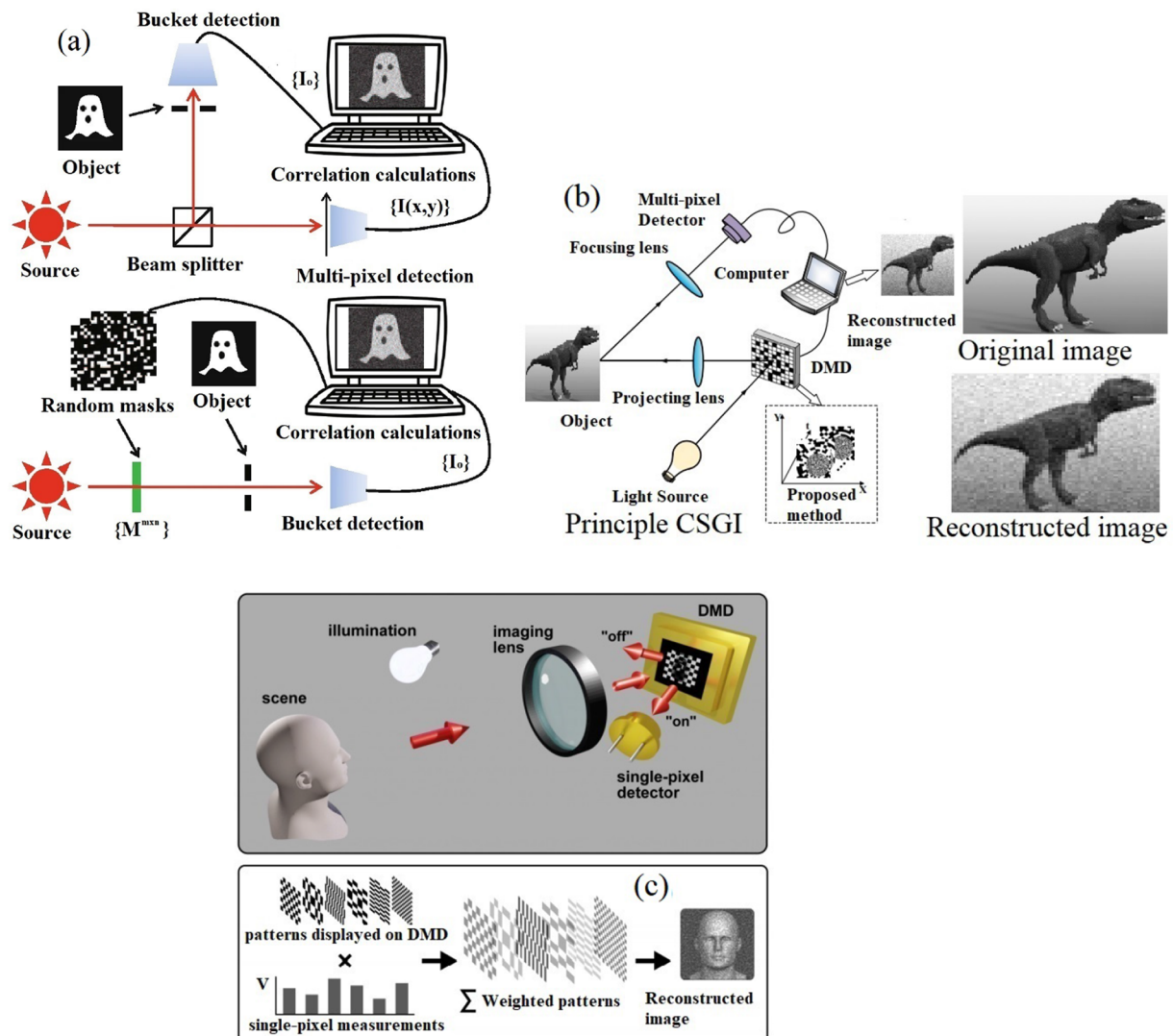
Here,  $I(x, y)$  represent the raw image. CSGI allows for using arbitrary structured patterns for spatial light modulation, such as random patterns, Hadamard basis patterns, Fourier basis patterns, discrete cosine basis patterns, or wavelet basis patterns [depicted in Fig. 6(b)]. To reconstruct the object image, CSGI attempts to search for a candidate of  $I(x, y)$ , which can minimize  $G$  (4). Different algorithms have been proposed for this task, such as L1-magic,<sup>48</sup> orthogonal matching,<sup>45</sup> augmented Lagrangian, an alternating direction algorithm (TVAL3),<sup>42</sup> or NESTA<sup>43,44</sup> algorithm (see Tables II and V). To date, many improvements have been made to the basic GI algorithm to be able to recover images more efficiently,<sup>28,44</sup> among which the following can be found: differential GI,<sup>44</sup> normalized GI (NGI),<sup>49</sup> corresponding imaging,<sup>50</sup> pseudoreverse GI,<sup>51</sup> sinusoidal GI,<sup>52</sup> and GI using the Hadamard basis.<sup>28</sup>

### C. Hadamard single-pixel imaging (HSI)

The Hadamard pattern<sup>53</sup> Eq. (5) is the most commonly used one in reconstructions of SPI images, mainly due to its orthogonality properties [depicted in Fig. 6(c)]. For generating a Hadamard matrix, a square matrix is first defined where its components are +1 or -1 with an agreement of two distinct rows in exactly  $n/2$  positions (and thus a disagreement in exactly  $n/2$  positions). This generated matrix  $H$  should satisfy the condition of  $HHT = nI$ , where  $T$  is the transposition of the matrix  $H$ ,  $I$  stands for an identity matrix, and  $N$  is the order of the matrix and can be generated by the Kronecker product expressed through Eq. (5), where 2 is an integer smaller than  $k$  and the matrix size is  $M \times N$  for ( $M = N$ ), as expressed by Eq. (6),<sup>53</sup> with  $m = 1, 2, 3, \dots, M$  and  $n = 1, 2, 3, \dots, N$ ,

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix} = H_2 \otimes H_{2^{k-1}}, \quad (5)$$

$$H_{2^k} = \begin{bmatrix} H(1, 1) & H(1, 2) & \dots & H(1, N) \\ H(2, 1) & H(2, 2) & \dots & H(2, N) \\ \dots & \dots & \dots & \dots \\ H(M, 1) & H(M, 2) & \dots & H(M, N) \end{bmatrix}. \quad (6)$$



**FIG. 6.** Overview of the sampling and scanning strategy used for single-pixel imaging reconstruction: (a) Computational Ghost Imaging (CGI) [Reprinted with permission from H.-C. Liu, *Sci. Rep.* **10**, 14626 (2020). Copyright 2020 Author(s), licensed under a Creative Commons Attribution 4.0 License]. (b) Compressive Sensing Ghost Imaging (CSGI), adapted with permission from<sup>47</sup> Cao *et al.*, *Sensors* **20**, 7093 (2020). Copyright 2020 Author(s), licensed under a Creative Commons Attribution 4.0 License. (c) Hadamard Single-Pixel Imaging (HSI) [Reprinted with permission from G. M. Gibson, S. D. Johnson, and M. J. Padgett, *Opt. Express* **28**, 28190–28208 (2020). Copyright 2020 Author(s), licensed under a Creative Commons Attribution 4.0 License].

To construct Hadamard's sequence (see Fig. 7), Sylvester's principle of recursive generation matrix was used, as expressed by Eq. (6), to generate the Hadamard matrix  $H_2^k(m, n)$ .

There are different orders for Hadamard Matrix generation as Cake-Cutting (CC),<sup>21</sup> Russian Doll (RD),<sup>54</sup> and Origami (ORCS).<sup>55</sup> For CC, the Hadamard matrix rows are organized so that the connected components follow a rising order. Each connected component corresponds to a group of pixels with the same value [depicted in Fig. 8(a)]. The RD is a scaling order method by a factor 2 of the Hadamard matrix [depicted in Fig. 8(b)]. This method provides a better signal-to-noise ratio (SNR) of the reconstructed image SPI.

Origami is a method of manipulation of patterns Hadamard that for generation needs to divide basic patterns them into  $n/4$  groups, where each new pattern will be generated to change the sequence of "0" and "1," in the direction vertical and horizontal [depicted in Fig. 8(c)].

#### D. Fourier single-pixel imaging (FSI)

FSI acquires the Fourier transform<sup>56</sup> of the object image and reconstructs the object image by applying an inverse Fourier transform. The Fourier transform of an image is a complete set of Fourier

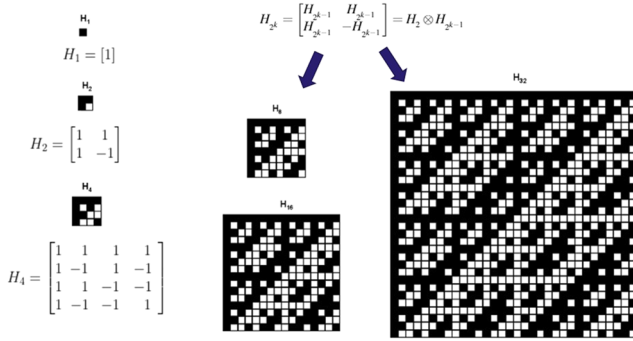


FIG. 7. Example of Hadamard matrix generation  $H_1, H_2, H_4, H_8, H_{16},$  and  $H_{32}$  using Sylvester methods.

coefficients [depicted in Fig. 9(a)]. Each coefficient is the weight corresponding to a unique Fourier basis pattern (also known as a “sinusoidal pattern” or “fringe pattern”). A Fourier basis pattern  $P_\varphi(x, y)$  is obtained by applying an inverse Fourier transform to a Dirac delta function  $\delta F(u, v, \varphi)$  as expressed in the following equation:<sup>56</sup>

$$P_\varphi(x, y) = \frac{1}{2} [1 + \text{real}\{F^{-1}\{\delta_F(u, v, \varphi)\}\}]. \quad (7)$$

Here,  $(u, v)$  is the coordinate in the Fourier domain (or spatial-frequency domain),  $\varphi$  is the initial phase, *real* denotes the real part,  $F^{-1}$  denotes an inverse Fourier transform, and  $\delta_F(u, v\varphi)$  is defined as shown in the following equation:<sup>56</sup>

$$\delta_F(u, v, \varphi) = \begin{cases} e^{j\varphi}, & u = u_0, v = v_0 \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

### E. Binary Fourier single-pixel imaging (BFSI)

BFSI can be implemented following two methods: (1) the four-step phase-shifting FSI and (2) the three-step phase-shifting

FSI [depicted in Fig. 9(b)]. In both methods, each Fourier coefficient  $F(u, v)$  can be acquired by differential measurements performing 4 and 3 measurements, respectively.<sup>59</sup> The four-step phase-shifting FSI allows for acquiring each complex-valued Fourier coefficient  $F(u, v)$  expressed by Eq. (9),<sup>59</sup> by using the following four patterns:  $P_0(x, y), P_{\pi/2}(x, y), P_\pi(x, y),$  and  $P_{3\pi/2}(x, y)$  during the corresponding four single-pixel measurements:  $D_0, D_{\pi/2}, D_\pi,$  and  $D_{3\pi/2},$

$$F(u, v) = (D_\pi - D_0) + (D_{3\pi/2} - D_{\pi/2}). \quad (9)$$

In the case of the three-step phase-shifting, FSI acquires each complex-valued Fourier coefficient with three single-pixel measurements  $(u, v)$  and applying the following equation:<sup>59</sup>

$$F(u, v) = (2D_0 - D_{2\pi/3} - D_{4\pi/3}) + \sqrt{3}j(2D_{2\pi/3} - D_{4\pi/3}). \quad (10)$$

Here,  $D_0, D_{2\pi/3},$  and  $D_{4\pi/3}$  are the single-pixel measurements corresponding to each applied structured pattern:  $P_0(x, y), P_{2\pi/3}(x, y),$  and  $P_{4\pi/3}(x, y).$  In both implementations of the FSI, the number of Fourier coefficients is the same as that of image pixels  $(M \times N)$  obtained at the end. Therefore, it is possible to acquire the coefficients in a specific order by using different type sampling paths that can be circular, diamond-shaped, or spiral,<sup>59</sup> where more important coefficients get higher priority. The four-step phase-shifting FSI and the three-step phase-shifting FSI are, respectively, essentially differential measurement methods, one symmetrically and one asymmetrically repetitive. For the image reconstruction in this case, it is necessary to apply a 2D inverse Fourier transform to the SPD output signal matrix. As Fourier basis patterns are initially grayscale, FSI has difficulty in taking advantage of high-speed binary pattern generation allowed by DMDs. Hence, Fourier basis pattern binarization using the up-sample-and-dither strategy is required in this case to reduce the quantization errors induced through binarization. The patterns to be binarized should be up-sampled with  $k > 1.$  To perform the up-sampling, interpolation of the image with the  $K^2$ -fold number of patterns equivalent to the application of the binarization algorithm based on error diffusion is

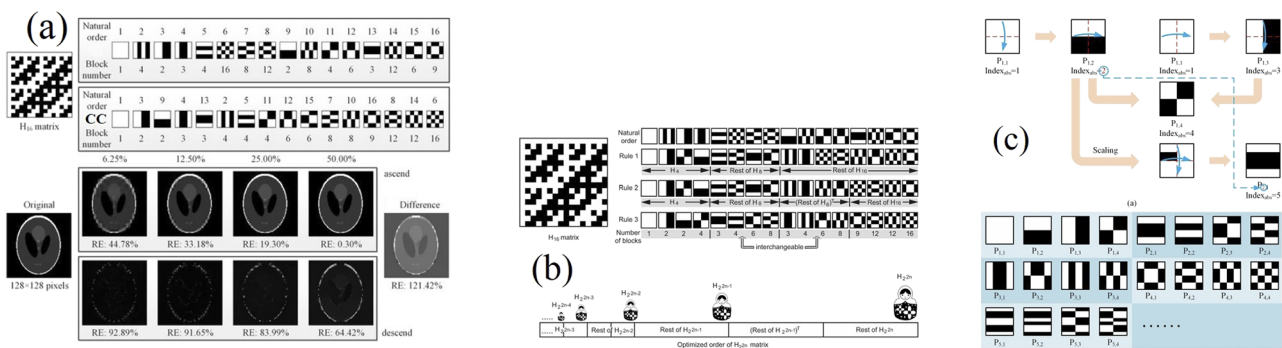
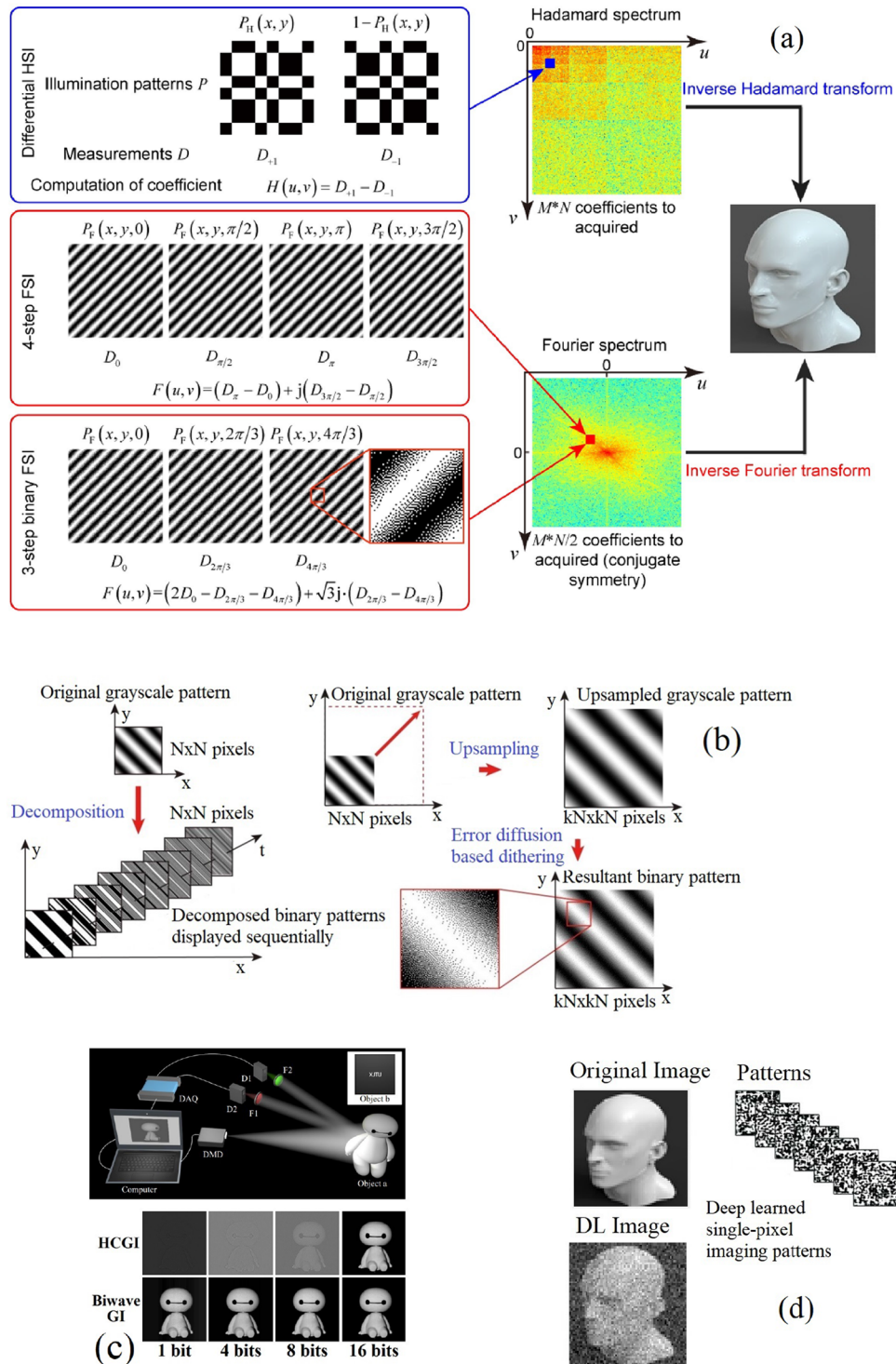


FIG. 8. Hadamard patterns generation CC, RD, and ORCS: (a) cake-cutting (CC)<sup>21</sup> Hadamard basis [Reprinted with permission from W.-K. Yu, *Sensors* **19**, 4122 (2019). Copyright 2019 Author(s), licensed under a Creative Commons Attribution 4.0 License]. (b) Russian Dolls (RD) Hadamard ordering [Reprinted with permission from Sun *et al.*, *Sci. Rep.* **7**, 3464 (2017). Copyright 2017 Author(s), licensed under a Creative Commons Attribution 4.0 License]. (c) Origami pattern construction (ORCS) [Reprinted with permission from W.-K. Yu and Y.-M. Liu, *Sensors* **19**, 5135 (2019). Copyright 2019 Author(s), licensed under a Creative Commons Attribution 4.0 License].





**FIG. 9.** Overview of the sampling and scanning strategy used for single-pixel imaging reconstruction: (a) FSI, adapted with permission from Zhang *et al.*, *Opt. Express* **25**, 19619–19639 (2017). Copyright 2017 Author(s), licensed under a Creative Commons Attribution 4.0 License. (b) BFSI<sup>57</sup> [Reprinted with permission from Zhang *et al.*, *Sci. Rep.* **7**, 12029 (2017). Copyright 2017 Author(s), licensed under a Creative Commons Attribution 4.0 License]. (c) WT [Reprinted with permission from Xi *et al.*, *Opt. Express* **27**, 32349–32359 (2019). Copyright 2019 Author(s), licensed under a Creative Commons Attribution 4.0 License]. (d) ML, adapted with permission from Gibson *et al.*, *Opt. Express* **28**, 28190–28208 (2020). Copyright 2020 Author(s), licensed under a Creative Commons Attribution 4.0 License.

required. If  $k = 2$  is used, it can effectively suppress the quantization errors.<sup>56,59</sup>

### F. Wavelet (WT) single-pixel imaging

Wavelets are mathematical functions<sup>59</sup> that map data onto different frequency components, where each component has a scale resolution [depicted in Fig. 9(c)]. WT has advantages over the Fourier method under condition discontinuities.<sup>60</sup> In application, SPI, the Haar wavelet,<sup>60</sup> is selected to be the simplest wavelet, the mother wavelet function of which is a binary function expressed in Eq. (11) and a 2D matrix defined in Eqs. (12) and (13).<sup>58,61</sup>

For implementation, due to the fact that Haar wavelets consist of a sequence of +1, 0, and -1, it is necessary to use two light frequencies to represent +1 and -1, respectively. The image is then reconstructed by inverse wavelet transform as expressed by the following equation:

$$\varphi(t) = \begin{cases} 1, & t \in [0, 1/2] \\ -1, & t \in [1/2, 1] \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

$$M_j(x, y) = \sqrt{2^{s-q}} \varphi(2^{s-q}((y-1)n+x) - k), \quad (12)$$

$$I = M_j^{-1} B. \quad (13)$$

### G. Machine learning (ML) single-pixel imaging

Applying the machine learning technique to SPI is the most recent approach that uses deep learning in a convolutional neural network (CNN) to reconstruct images based on fewer measurements than those required in other methods, such as orthogonal sampling or the ghost imaging technique.<sup>62</sup> The integration of Graphics Processing Units (GPUs) for CNN allows higher computation rates than those achieved by conventional computer processors. Hence, applying a CNN together with the computational GI approach allows us to reconstruct an image using a minimal number of measurements.<sup>62</sup> In some applications, SPI does not need to perform a full image reconstruction to detect and classify objects, so it can identify even very fast-moving objects due to this advantage.<sup>29,62,63</sup> Applying a CNN [depicted in Fig. 9(d)] for image reconstruction with fewer samples could be adapted into a control system for an autonomous vehicle using image-free classification sensing schemes.<sup>62,64</sup>

### H. Discussion on different single-pixel imaging approaches

Regarding SPI image reconstruction methods, HSI and ML have proved to be more adaptable to existing illumination modulation technologies using DMD devices or even LED arrays in its simpler form. The ML is a good candidate that already allows for doing image reconstruction with a high compression factor without the need to use  $N \times N$  patterns, which allows the user to decrease the SPD output signal processing times while still attaining quite acceptable image quality.<sup>48</sup> It is an ideal method for applications requiring video streaming or forming part of navigation systems where a quick response, i.e., very low data processing time for each

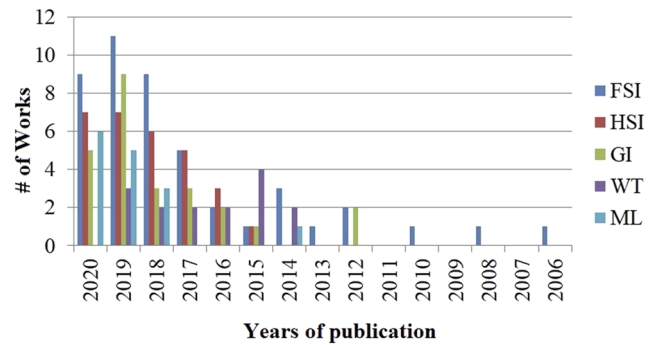


FIG. 10. The number of works reported applied in SPI using the methods from CGI, HSI, FSI, WT, and ML from 2006 to 2020, FSI with 46 published, HSI with 29 published, GI with 25 published, WL with 15 published, and ML with 15 published.

individual image, is needed. ML can be combined with algorithms developed for the CSGI approach<sup>3,49,50</sup> and can adapt deep learning to train a sequence of projected patterns or improve the neural activation network used in the CNN process. ML implementations require GPUs and require large databases for network training. Hence, this approach does have limitations under certain conditions, noise or untrained networks. The FSI<sup>56</sup> and WT<sup>59</sup> methods apply the inverse transforms for the reconstruction of SPI images. Therefore, the application of FFT and FWT algorithms on GPU is possible for optimization of the rebuild time and the possibility of using only 3%–10% of the projected patterns. Unfortunately, these methods cannot be integrated directly in systems using devices such as DMDs, so it might be necessary to apply quantization techniques, smoothing, and up-sampling strategies for binarization of the patterns producing quantization noise when using them, all of which might increase the processing times required and degrade the quality of the reconstructed images.<sup>18,65</sup> The CGI-based techniques use randomly generated illumination patterns that have the disadvantage of, unlike Hadamard that has properties of orthogonality, requiring a maximum amount of  $N \times N$  patterns. However, in the case of random patterns, it would be necessary to generate many patterns to enable for the reconstruction of an image with good quality, which will inevitably increase the data processing time required to generate an image. One solution is to use CSGI- or ML-based techniques. The GI is adaptable to SPI systems using technologies such as DMD, LED, SLM, and pseudothermal light source, which gives the greater capacity for rebuilding applications of images in the x-ray<sup>66</sup> and THz<sup>67,68</sup> spectra. In recent years, an increase in the use of the FSI method can be observed (see Fig. 10), caused by the application of FSI improved algorithms that reduce low-frequency sampling, maintaining the image quality by using random sampling techniques.<sup>69</sup> Some methods, such as ML, have been recently implemented in several SPI systems, mainly due to advances in the use of GPU architectures, normally used for fast-moving object detection.<sup>29,62</sup>

## III. MODULATION TECHNOLOGIES

As previously shown in Table II, there are several options regarding the modulation technologies used to produce the

patterns for either structured detection or structured illumination in SPI, among which the following can be found: Liquid Crystal Spatial Light Modulators (LC-SLMs), Digital Micromirror Devices (DMDs), LED arrays, or pseudothermal light source.<sup>38</sup>

### A. Liquid crystal spatial light modulators (LC-SLMs)

A liquid crystal spatial light modulator (LC-SLM)<sup>38,70</sup> is an electrically programmable device that modulates light according to a fixed spatial (pixel) pattern induced. SLM is typically used to control incident light in amplitude, phase, or through the combination of both. When a polarized light beam enters the SLM device, it passes through an array of glass structures with transparent electrodes and liquid crystal layers. If a certain voltage is induced on the electrodes, an electric field is produced that causes a change in the LC layer's optical properties and generates a phase pattern in each pixel, which causes the light to be reflected or not, a technology which can be combined with GI schemes using LC-SLM as a single-pixel detector<sup>32</sup> (see Table III).

### B. Digital micromirror devices (DMD)

Digital Micromirror Devices (DMDs),<sup>38</sup> consisting of an array of hundreds of thousands of individually addressable micromirrors, offer a method of modulating light that is fast and works over a broad range of wavelengths. Micromirrors can be individually oriented at  $\pm 12^\circ$ , to the array's plane, by displaying a binary pattern on the DMD if illuminated using a structured light pattern. The DMDs are commercially available, having binary pattern display rates of below 40 kHz, which allows for near-video rate image reconstruction on a standard performance computer for relatively low-resolution applications<sup>60</sup> (see Table III).

### C. LED arrays

The LED arrays<sup>38,71</sup> can offer a solution to capture images of a dynamic scene with a rate higher than the one normally achieved by DMDs using high-speed structured illumination with a switching time of the LEDs below 1  $\mu$ s. This type of technology can adapt to Hadamard pattern generation and compressive sensing algorithms used for image reconstruction based on the SPI<sup>37</sup> principle (see Table III).

TABLE III. Summary of modulation methods.<sup>38</sup>

Technology	Advantage	Disadvantage
LC-SLM <sup>38,70</sup>	Grayscale modulation and programmable	Slow modulation and low-power endurance
DMD <sup>38,74</sup>	Faster than LC-SLM and programmable	Binary modulation and not fast enough
LED array <sup>38,71</sup>	Much faster than DMD and programmable	Binary modulation and structured illumination only random
Pseudo-thermal <sup>38,72</sup>	Much faster than DMD and controller	Modulation and complicated manufacturing

### D. Pseudo thermal light source

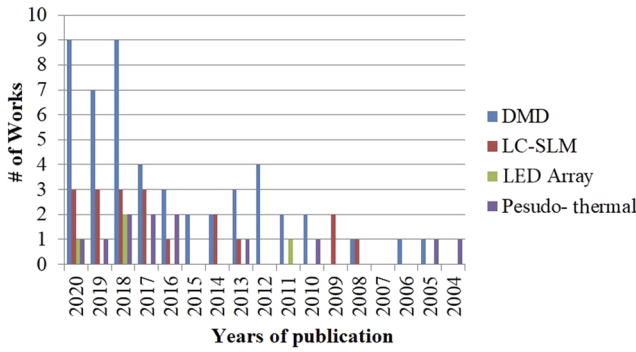
Another option for generating illumination patterns in the GI scheme is the use of pseudo-thermal light sources, which consist in applying a laser beam through a rotating ground-glass diffuser.<sup>72</sup> The rotary action of the diffuser causes the cross-sectional intensity of the resulting optical beam to vary with time, creating two optical beam splitter forms, i.e., 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 pseudo-thermal source are determined by the properties of the materials from which it is made, and the light that emerges of a pseudo-thermal source is compared for its coherence properties to the light originated in a LED source<sup>73</sup> (see Table III).

### E. Discussion on different modulation technologies

For SPI applications, DMD as a projection element has become the most successful candidate for different scanning methods (HSI, FSI, GI, WT, and ML). In recent years, DMDs have increased their modulation frequency from 20 to 40 KHz<sup>74,75</sup> but still remain quite slow if compared to commercially available LED arrays.<sup>37,59</sup> LED arrays allow lower costs in applications with required operational speeds that must be higher than those used by other technologies. For applications requiring video streaming or 3D image reconstruction,<sup>76</sup> high processing speeds have been achieved,<sup>59,77</sup> although having relatively low spatial resolutions of, e.g.,  $32 \times 32$  pixels<sup>37</sup> and also some limitations regarding the incident light power at 5 mW.<sup>77</sup> Pseudothermal light source technology has become the element of choice for applications in the THz spectrum<sup>68</sup> with better accuracy and robustness to noise if compared to DMDs.<sup>69</sup> In terms of the work published in recent years regarding the SPI, an increase in using machine learning methods can be observed (see Fig. 10). The trend indicates that in successive years, ML will displace other methods, such as FSI and HIS. The spatial light modulator (LC-SLM) has many applications in SPI if we compare it with the DMD technology in what its frequency resolution is concerned. The DMD is faster than SLM and also less expensive. Following the improvements in the liquid crystal display (LCD) technology, it could be used as an intensity-only SLM to conduct structured illumination, which would allow SPI to be performed in a lens-less way.<sup>78</sup> Nevertheless, the DMD technology still has many applications in the SPI (see Fig. 11) domain. Although there have been improvements in modulation frequency, DMDs still have frequency limitations for some applications, such as video streaming and 3D imaging. The LED technology applied to SPI is still in the testing phase. It is possible that in the future, it could displace the DMD technology due to its low-cost<sup>37</sup> and high modulation frequencies.

## IV. SINGLE-PIXEL RECONSTRUCTION ALGORITHMS

In the past few years, different algorithms have been proposed for SPI reconstruction in the literature (see Fig. 5), which are classified into three categories according to their iteration type (see Table IV): the non-iterative methods (DGI), the linear iterative methods (GD, CGD, Poisson maximum likelihood method, and AP), and the nonlinear iterative methods (sparse representation method, TV).



**FIG. 11.** Yearly analysis of the amount of publications regarding the existing technologies used in SPI between 2006 and 2020. As it can be observed in the graph, DMD appears in 50 publications, LC-SLM in 19, SPI using LED arrays in 4, and pseudothermal light source in 12 publications. Note that DMD technology is the mostly used technology in SPI applications if compared with other technologies, such as those using LED arrays for structured illumination or other similar technologies.

- Non-iterative methods (see Algorithm 1) perform direct reconstruction of the linear representation system SPI  $Ax = b$  without iteration (where  $A \in \mathbb{R}^{m \times n}$  denote the light modulation matrix  $m$  patterns that represent  $n$  pixels,  $x \in$

$\mathbb{R}^{n \times 1}$  denote the target scene, and  $b \in \mathbb{R}^{m \times 1}$  is the measurement vector) through the calculation  $x = (A^T A)^{-1} A^T b$  based on the fact that SPI measurement stands for the correlation between modulation patterns and the target scene  $x$ . These methods are based on linear correlation, requiring a more significant measurement value to approximate the target scene. More measurements allow for producing a high-quality reconstruction. Some linear correlation methods have been proposed as ghost imaging (DGI) (see Algorithm 1),<sup>28</sup> takes consideration illumination fluctuations the pattern intensity  $s \in \mathbb{R}^{m \times 1}$  (14),<sup>28</sup> with  $\{b_i\} = \frac{1}{m} \sum_{i=1}^m b_i$  as measurement average and  $\{a_i\} = \frac{1}{m} \sum_{i=1}^m a_i$  as modulation pattern (row) intensity average, to reconstruction SPI (some strategies SPI as CSGI and CGI used these types of algorithms),

$$x = \frac{\{b_i\}}{\{s_i\}} \{a_i\}. \quad (14)$$

- Lineal iterative methods are based on solving the quadratic minimization problem  $\min \|Ax - b\|_2^2$  through the gradient descent (GD)<sup>40</sup> or conjugate gradient descent (CGD)<sup>79</sup> methods for the reconstruction of an  $n$ -pixel SPI image. Other methods use photons signal statistics to arrive at sensors and spectral properties of the light field as the alternating projection (AP) method.<sup>80</sup> Due to the photons

**TABLE IV.** Comparison of different SPI reconstruction algorithms.<sup>40</sup>

Algorithm	Method	Principles
No iteration	Matrix inversion formation conventional correlation <sup>27</sup> and DGI <sup>28</sup>	Formation fitting measurement: scene-pattern correlation $x = (A^T A)^{-1} A^T b$
Linear iteration		Problem
	Gradient descent (GD) <sup>40</sup>	Formation fitting $\min \ Ax - b\ _2^2$
	Conjugate gradient descent (CGD) <sup>79</sup>	$A^T Ax = A^T b$
	Poisson maximum likelihood <sup>82</sup>	$\sum_{i=1}^m (a_i x - b_i \log(a_i x))$
	Alternating projection (AP) <sup>40</sup>	Measurement: zero-spatial frequency coefficient $Ax = b$
		Gradient p
		$A^T (Ax - b)$
		$-r^{k-1} - \frac{r^{(k-1)r} r^{(k-1)}}{r^{(k-2)r} r^{(k-2)}} \mathcal{P}$
		$A^T \left( \frac{Ax-b}{Ax} \right)$
		$\frac{a_i^T \odot (a_i^T \odot x)}{\max(a_i)^2} \frac{a_i x - b_i}{a_i x}$
		Step $\Delta x$
		$-\frac{p^T A^T r}{p^T A^T A p}$
		$-\frac{r^T r}{p^T A^T A p^k}$
		Algorithm 2
		1
Non-linear iteration	Sparse representation (CS) <sup>5,85</sup>	Image prior $proj = A^T r \operatorname{argmax} \ proj\ $
	Total variation (TV) <sup>83</sup>	$Dx = c, Ax = b$
		$\left( (A^T A)^{-1} A^T \right) b$
		$\ y - Ax\ _2 < \epsilon$
		$\min L = \ c\ _1 + \frac{\mu_1}{2} \ Dx - c\ _2^2 + \frac{\mu_2}{2} \ Ax - b\ _2^2 + \frac{\mu_2}{2} \ b\ _2^2$

**ALGORITHM 1.** Differential ghost imaging (DGI) SPI.<sup>40</sup>

---

**Input :** signal  $y$ , pattern  
**Output:** target scene  $x$   
 initialize variables  $\bar{y} = 0, \bar{y}_i = 0, \bar{r} = 0, \bar{r}_i = 0, iter = 0$   
**while do**  
      $iter = iter + 1$   $\bar{y}_i = (\bar{y} * (iter - 1) + pattern * y) / iter$   
      $\bar{y} = (\bar{y} * (iter - 1) + y) / iter$   
      $\bar{r} = (\bar{r} * (iter - 1) + \sum_{i=0}^M \sum_{j=0}^N (pattern_{MN})) / iter$   
      $\bar{r}_i =$   
      $(\bar{r}_i * (iter - 1) + \sum_{i=0}^M \sum_{j=0}^N (pattern_{MN})) * pattern / iter$   
      $x = \bar{y}_i - (y / \bar{r}) * \bar{r}_i$   
 $x = x - \Delta x * p$  **return**  $x$

---

following Poisson distribution,<sup>81</sup> the SPI reconstruction can proceed through the maximum likelihood estimation method<sup>82</sup> (see Algorithm 2). This method aims at estimating  $x$  by maximizing the likelihood the measurements  $b_i \in b$  (15),<sup>82</sup>

$$\min L(x) = \sum_{i=1}^m (a_i x - b_i \log(a_i x)). \quad (15)$$

The AP is a SPI reconstruction method from the view of the spatial spectrum, the signal light arrived to the photodiode, where measurement  $b_i \in b$  is related to the light field's zero-spatial-frequency coefficient. In spatial space, the target image  $x$  can be represented<sup>80</sup> by

$$x' = x - \frac{a_i^T}{\max(a_i)^2} \frac{a_i x - b_i}{n}. \quad (16)$$

- Nonlinear iterative methods aim at SPI reconstruction using a reduced number of measurements. There are two methods that are widely used, including the spatial representation and the total variation (TV).<sup>83</sup> The first method is based on the theory of compressive sensing (CS)<sup>5</sup> that seeks to solve the least-squares problem  $\|y - Ax\|_{l_2}^2$  through using an iterative algorithm that seeks the maximum correlations between the measurements and the matrix  $A$ , called dictionary, which contains in each column atom. First, determine the target scene  $x$ . Then, in each iteration type, algorithms seek to select the best atom iteratively until we can reduce the approximating error or reach a sure accuracy (see Algorithm 3). Once we reach the optimal atom, we can get the signal recovered  $x$ . The CS methods more used for SPI reconstruction as orthogonal matching,<sup>45</sup> OMP,<sup>84</sup>

**ALGORITHM 2.** Poisson SPI.<sup>40</sup>

---

**Input :** Procedure Input signal  $y$   
 $p, \alpha \in [0.01, 0.3], \beta \in [0.1, 0.8]$   
**Output:** target scene  $x$   
 calculate gradient  $p$   
 $\Delta x \leftarrow 1$   
**while**  $L(x + \Delta x p) > L(x) - \alpha \Delta x^T p$  **do**  
      $\Delta x \leftarrow \beta \Delta x$   
 $x = x - \Delta x * p$  **return**  $x$

---

**ALGORITHM 3.** Orthogonal matching pursuit OMP.<sup>85</sup>

---

**Input :** Dictionary  $A$ , signal  $y$ , target sparsity  $K$   
**Output:** Sparse representation  $x$  such that  $\bar{y} = Ax$   
 set  $I = \{\emptyset\}, r = x, A = 0$   
**while** *stopping test* **false do**  
      $proj = A^T r$  Matching atoms with residual  
      $k = \operatorname{argmax}_k |proj|$  Finding the new atom best match  
     with residual  
      $I = (I, k)$  Support update  
      $x = ((A^T A)^{-1} A^T) y$  Calculate least squares  
     pseudoinverse  
      $\bar{y} = Ax$  Approximation update  
      $r = \bar{y} - y$  Calculate residual  
      $\|y - Ax\|_2 \leq \epsilon$  Calculate target error  
**return**  $\bar{y}$

---

batch-OMP,<sup>85</sup> fast-OMP,<sup>86</sup> and others can be integrated into FPGA or GPU technologies.

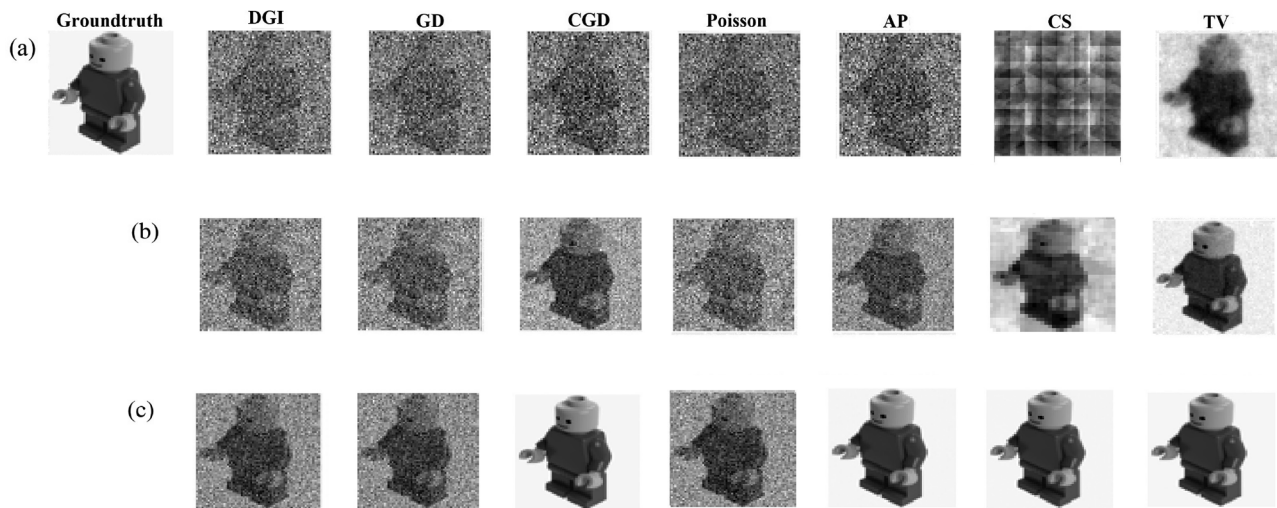
The TV method is based on the calculation of the gradient for SPI reconstruction. Through minimizing the Lagrangian function (17), we can determine the target scene  $x$  (18). The (TVL3)<sup>42</sup> and NESTA<sup>43,44</sup> algorithms use this type of method,

$$\min L = \|c\|_{l_1} + \frac{\mu_1}{2} \|Dx - c + \frac{y_1}{\mu_1}\|_{l_2}^2 + \frac{\mu_2}{2} \|Ax - b + \frac{y_2}{\mu_2}\|_{l_2}^2, \quad (17)$$

$$x = (\mu_1 D^T D + \mu_2 A^T A)^{-1} \left[ \mu_1 D^T \left( c - \frac{y_1}{\mu_1} \right) + \mu_2 D^T \left( c - \frac{y_2}{\mu_2} \right) \right]. \quad (18)$$

**A. Discussion on different reconstruction algorithms: single-pixel**

For SPI reconstruction, defining that algorithm is the best is a factor that depends on the application development. In the case of reconstruction with small-scale images, the DGI, TV [depicted in Fig. 12(a)], and CS [depicted in Fig. 12(b)] methods need the least measurements and minimum processing time (see Fig. 5). This advantage allows their application in vision close to real-time. At the same time, Poisson and GD methods need the highest number of measurements and processing time higher in comparison with other methods [depicted in Fig. 12(c)]. For the case, large-scale images, the CGD, and AP methods run fastest (see Fig. 5). In noise conditions, the TV, CS, and AP methods are the most robust but are need to take into consideration that there must be a trade-off between capture efficiency, computational complexity, and noise robustness among which we like to choose which is the most efficient method to the SPI image reconstruction. At the level of memory and computation complexity, the DGI and AP methods are the simplest to implement, and for being non-global, methods do not need to store all the patterns during the processing of calculation, which gives them a high storage efficiency advantage, for example, its implementation into device FPGA.



**FIG. 12.** SPI simulation results of different algorithms under different sampling ratios: (a) sampling ratio = 0.2, (b) sampling ratio = 0.8, and (c) sampling ratio = 3. We can see that CS and TV methods need few samples for SPI reconstruction, while DGI, GD, and Poisson need a higher sampling ratio  $>1$ .

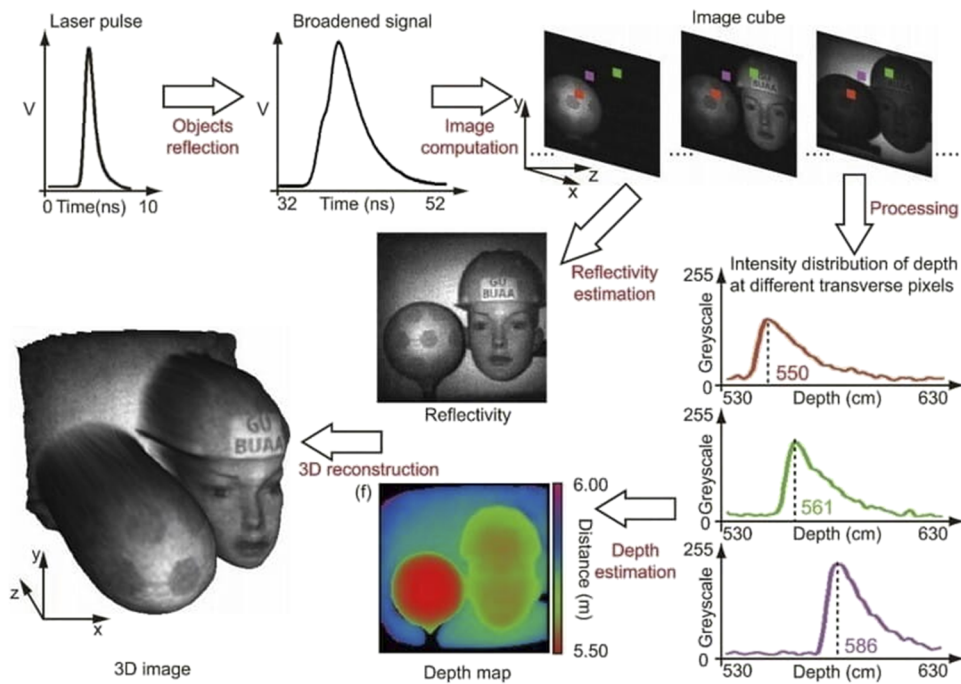
## V. SINGLE-PIXEL 3D IMAGING AND RANGING

3D imaging and ranging represents an intensively explored technique with a wide range of applications, including object detection, surface mapping, and 3D scenario mapping for autonomous vehicles. Within the field of 3D imaging, there are different principles used that include time-of-flight,<sup>87,88</sup> stereo vision,<sup>42,89,90</sup> and 3D computational ghost imaging<sup>7,26</sup> approaches, each having their advantages and drawbacks, which depend on the specific application for which they are developed.

### A. Time-of-flight imaging

The Time-of-Flight (ToF) measurement method determines the distance  $d$  of a scene by actively illuminating it with pulsed light (normally emitted by lasers) or continuous-wave modulated light (normally emitted by an array of LEDs) and comparing the detection time  $t_a$  of the back-scattered light to the time of the illumination pulse  $t_0$  (i.e.,  $d = \Delta t c/2$ ), where  $\Delta t = (t_a - t_0)$  is the time of flight and  $c$  is the speed of light. Using in this way the obtained distance information of the objects existent in the illuminated scene, a 3D image of the surroundings of ToF systems can be obtained by combining a depth map (i.e., the 2D array containing the data about the distance of different objects in the scene to the photodetector) with a transverse reflectivity image of the same scene obtained by applying the SPI principle.<sup>88,91</sup> For ToF measurements, pulsed illumination using nanosecond pulsed lasers is the best option if combined with really fast photodetectors capable of near single-photon detection (such as single-photon avalanche diodes—SPADs, for example) due to temporal resolutions achieved in these systems the tens of picoseconds. Therefore, the ToF method is compatible with long-range, high-precision depth mapping. Some applications using this approach, capable of near single-photon counting,<sup>11,87</sup> offer the advantage of working with low light intensity: a very important asset if the active

illumination light reflected from objects placed at large distances is to be detected. One crucial problem in these applications is the signal-to-noise ratio (SNR), as the system must be able to discriminate the very weak light signals originated by active illumination from the normally quite high background illumination signals causing most of the photon-shot noise. One additional problem that has to be taken into consideration here is the fact that silicon-based photodetectors have much higher QE in the visible part of the spectra (i.e., much higher ability to detect background radiation) than in the NIR part of the spectra, a preferred kind of active illumination radiation making it “invisible” to the human eye. However, it has reducing detection efficiency due to dead time (typically 10 s of nanoseconds) between successive measurements. The latter can be importantly improved by using narrow band filters in front of the photodetector arrays that allow only active illumination radiation to come through, eliminating the background illumination negative influence. In recent developments based on SPAD arrays and laser pulses in nanoseconds,<sup>91</sup> the depth accuracy of 3 mm at a detection range of 5 m has been reported. There exist two additional approaches, when it comes to ToF approaches: direct ToF (d ToF), where the system measures directly the time it takes for the laser emitted light pulse to reach an object in the illuminated scene, get reflected by it, and finally reach the photodetector array and indirect ToF (i-ToF), where hundreds or thousands of pulses are emitted, and the amount of photons detected by the photodetector array is statistically analyzed and compared to the amount of photons emitted within each pulse duration to be able to determine the ratio between the two and so, indirectly, calculate the mean time it took for the emitted pulses to reach the objects in the scene, be reflected by them, and finally get detected by the photodetector array and in this manner obtain the depth map of the illuminated scene. The ToF method, in each of its different approaches, is definitively a good candidate for long-distance 3D



**FIG. 13.** Time-of-flight imaging 3D [M.-J. Sun and J.-M. Zhang, *Sensors* **19**, 732 (2019). Copyright 2019 Author(s), licensed under a Creative Commons Attribution 4.0 License].

measurement and offers several advantages over other 3D imaging techniques, such as stereo vision or structured light 3D imaging<sup>89,92</sup> (see Fig. 13).

### B. Stereo vision

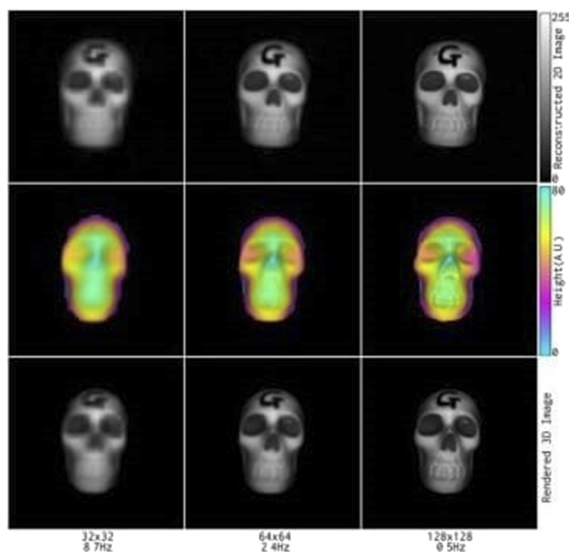
Stereo vision uses two or more images obtained simultaneously by using image sensors placed at different viewpoints in order to reconstruct a 3D image of the scene through the comparison of different geometries depicted in the images obtained. Another approach uses different illumination patterns for the simultaneous images taken (as in the photometric stereo<sup>41,90</sup> approach). To apply the stereo vision approach in SPI, it is necessary to detect the different projected light patterns that illuminate the objects in the depicted scene using multiple single-pixel detectors, placed at a separation of 90° from each other.<sup>76</sup> It is recommended to use the light pattern modulation generated by DMDs<sup>38</sup> or, in case the spatial resolution is not so important for the application of choice, a LED array,<sup>37,77</sup> an approach that can additionally improve the data processing speed for the 3D image reconstruction (see Fig. 14). Mainly, two factors define the performance of SPI-based stereo vision 3D image reconstruction: the quality of the 2D images obtained by the SPDs placed at different viewpoints and separated 90° from each other and the geometry of the system setup that determines the depth resolution that can be attained and that depends on the ratio of separation between the SPDs and their distance to the object to be depicted. Due to the limited depth resolution that can be obtained if stereo vision approach is followed, it can be employed in applications such as close industrial inspection or object 3D profiling (see Fig. 14).

### C. 3D computational ghost imaging (3D-CGI)

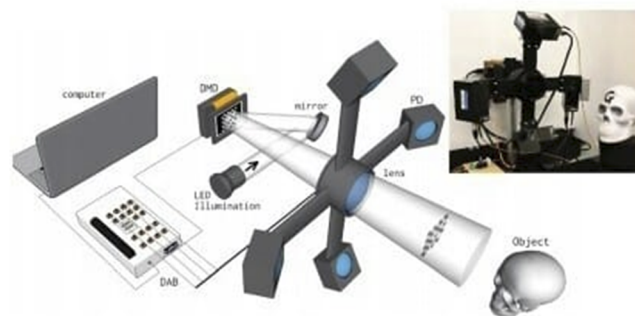
An alternative to the stereo vision technique, which needs that the scenario is aligned and has the correct geometry, could be applying the principle of photometric stereo.<sup>90</sup> The photometric stereo vision systems can capture a sequence of images where each one is obtained using different illumination setups. The resulting image sequence is much easier to align, provided that the sequence is captured fast enough to avoid movements of the objects in the scene between consecutive image frames. From the image sequence's shading profile, the surface normals can be estimated and used to finally generate the depth map of the illuminated scene. This technique is adaptable to computational GI<sup>26,93</sup> (see Fig. 15) using multiple single-pixel detectors rather than multiple illumination sources.<sup>7</sup> Placing multiple single-pixel detectors at different positions to depict the same illuminated scene allows for capturing multiple images using different illumination, similar to what is used in the photometric stereo vision systems: the shading in each simultaneously obtained SPI generated image appears as if it was illuminated from a different direction. The surface gradients could be estimated from the shading of SPI images and used to finally generate the depth map of the scene or the 3D reconstruction of the objects depicted<sup>90,94</sup> (see Fig. 15).

### D. Discussion on different single-pixel 3D imaging and ranging approaches

For 3D SPI applications, the ToF method allows for creating a high-precision depth mapping achieving resolutions of <10 mm at distances even larger 100 m. Due to the use of pulsed lasers,

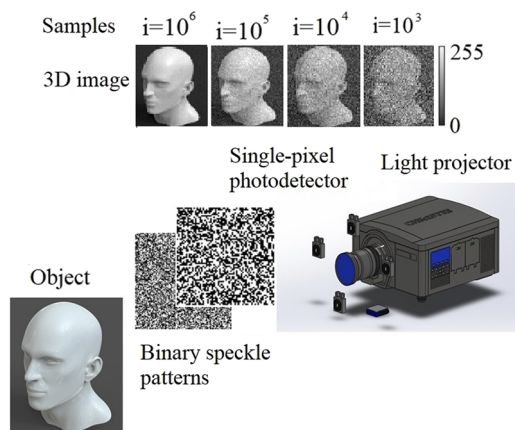


**FIG. 14.** Stereo vision 3D [Reprinted with permission from Zhang *et al.*, *J. Opt.* **18**, 035203 (2016). Copyright 2016 Author(s), licensed under a Creative Commons Attribution 4.0 License].

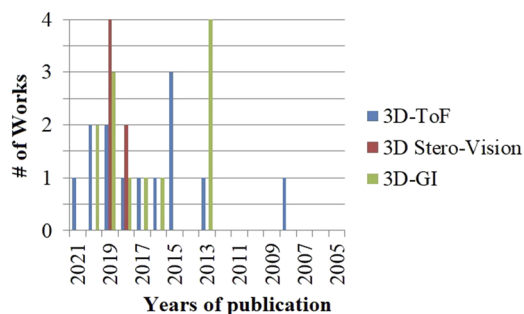


the time resolutions of order of the tens of picoseconds are achievable, enabling fast measurements. The limitations we detect for this approach lie mainly in the following: (1) the eye-safety regulation IEC62471 for class 3R lasers that limit the laser power to 5 mW<sup>95</sup> in the visible range that is also related to extremely low detection efficiencies of silicon-based photodetectors for active illumination with wavelengths of 850 nm or 905 nm usually used for this kind of applications, (2) the necessity to limit the background illumination, especially in the visible range, and (3) the fact that when interacting with particles with sizes in the micrometer or supra-micrometer range (as those present in fog, rain, snow, or smoke rich environments), several scattering mechanisms between the emitted photons and these particles take place, preventing the active

illumination radiation of propagating toward the objects in the scene or from these objects to the photodetectors trying to detect them.<sup>76</sup> For 3D SPI applications, SPAD detectors<sup>96</sup> can be used for applications such as the detection of moving objects using laser systems with pulse times of around 50 ps, motion detection in oblique environments, or recovery of three-dimensional objects behind the line of sight.<sup>97,98</sup> For 3D ToF applications, ML techniques are better suited for depth estimation and 3D reconstruction.<sup>96</sup> If the different stereo vision techniques applied to 3D SPI camera applications are to be compared to each other, the photometric stereo vision approach would be the most efficient. Hence, it would be advisable to use multiple (minimum 4) SPDs or, if not available, recover four SPI images obtained at different positions, even if not obtained simultaneously, in order to be able to perform the 3D image reconstruction using shape-from-shading (SFS) techniques<sup>90,94</sup> or deep learning.<sup>99</sup> These techniques can be merged with ToF to reach higher spatial



**FIG. 15.** 3D computational ghost imaging through the use of a photodetector array for this method of reconstruction is needed for a high number of samples.



**FIG. 16.** Number of works reported applied in SPI-3D use different methods as time-of-flight with 13 published, stereo vision with 6 published, and 3D computational ghost imaging with 12 published using the existing technologies such as DMD, LC-SLM, LED array, and pseudo-thermal.



resolutions of the final depth map. In terms of works published in recent years regarding the 3D SPI, the 3D ToF method (see Fig. 16) has proved to be a reliable method that offers quite acceptable spatial resolutions of <5 mm at 5 and 10 m distances for achieved time resolutions of between 50 and 100 ps, which can be adapted for 3D SPI used in static or moving scenes<sup>62</sup> having a wide field of applications, among others, also for autonomous navigation of unmanned vehicle. The other methods revised, used in 3D SPI generation, such as Vision Stereo, require multiple detectors and precise geometry adjustments, so they are not readily applicable for SPI. This might explain the limited number of publications reported on this approach. In contrast, the GI method combined with photometric stereo vision<sup>90</sup> is widely used in unconventional 3D applications, such as multi-wavelength imaging,<sup>13</sup> terahertz imaging,<sup>68</sup> and x-ray imaging.<sup>41</sup>

## VI. EVALUATION OF DIFFERENT SYSTEM ARCHITECTURES USED TO GENERATE SINGLE-PIXEL IMAGES

A limiting factor in single-pixel imaging systems using central processing units (CPUs) of conventional personal computers (PCs) is the total time required for signal collection and image reconstruction, taking ~30 min for a static image with spatial resolutions comparable to those produced by commercially available image sensors. The delays present in the data transport, inherent to the amount of data required to define the instruction cycles and generation of patterns sent to the DMD and the data collected from the SPD, the attainable data transfer speeds if USB protocols are used, and the computational cost of the algorithms used to reconstruct images, all limit the use of this hardware architecture for real-time SPI applications. An alternative is the integration of field-programmable gate-arrays (FPGAs) or embedded GPU devices. Elements such as FPGA are used to accelerate the hardware performance by managing memory accesses more efficiently using an architecture based on a pipeline to improve memory reading speeds<sup>100</sup> when processing the acquired data and reconstructing single-pixel images. Some other FPGA solutions use Synchronous Dynamic Random-Access Memories (SDRAMs) for data transport from the FPGA to the DMD, decreasing the time for overall signal collection.<sup>101</sup> It is necessary to consider the computational cost of the reconstruction algorithms used in SPI. One option is applying the orthogonal correspondence search (based, e.g., on the Orthogonal Matching Pursuit, OMP) algorithms<sup>85</sup> implemented for HSI or FSI. This algorithm is divided into two main stages: finding the closely correlated vectors and solving the least-squares problem by applying the conjugate gradient (CG) technique,<sup>100</sup> a calculation that unfortunately requires high processing times due to the fact that it requires inversion matrix calculations<sup>85</sup> or other similar methods, such as the one proposed by Quero.<sup>85</sup> Following this approach, some system solutions based on Xilinx FPGA (XC7VX690T),<sup>101</sup> Virtex-5,<sup>102</sup> and Virtex-6 FPGA<sup>103</sup> architectures have reported reconstruction times of 24  $\mu$ s for 128  $\times$  128 pixel image sizes (see Table V<sup>101,102</sup>).

The compressed sensing algorithms are very suited to be implemented in parallel on GPUs, an approach involving massive matrix/vector operations on the GPU platform to achieve good SPI performances. It is necessary to point out, nevertheless, that the bottlenecks, if the OMP algorithm is applied, lie in the projection

**TABLE V.** Comparison of different solutions of FPGA using OMP, applying to the reconstruction 2D image.

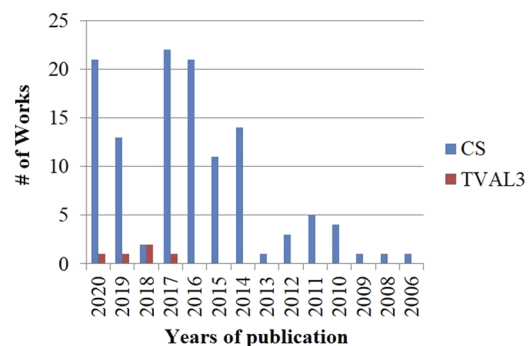
FPGA	Image size	Reconstruction time
Virtex-5 <sup>102</sup>	128 $\times$ 128	24 $\mu$ s
Virtex-6 <sup>103</sup>	128 $\times$ 1	16 $\mu$ s
Virtex-7 <sup>101</sup>	128 $\times$ 128	8.97 $\mu$ s

**TABLE VI.** Comparison of different SPI solutions based on both, FPGA and GPU platforms, using OMP<sup>103</sup> algorithms, applied for the reconstruction of 2D SPI images: the FPGA-based solutions yield processing times of 24  $\mu$ s, with an improvement of  $\times 2.67$ , while GPU-based architectures yield processing times in the order of 37.5 ms, an improvement of  $\times 3.4$ .

FPGA	Image size	Reconstruction time
Virtex-7	128 $\times$ 128	8.97 $\mu$ s <sup>103</sup>
n-Nvidia GeForce <sup>104</sup>	128 $\times$ 128	11 ms

module and the module responsible for calculation of the least-squares solutions. Some of the solutions proposed to speed up the processing task of the projection module, such as Fujimoto's matrix-vector multiplication algorithm, have been reported, as well as the matrix-inverse-update calculation, to speed up the least-squares calculation module.<sup>104</sup> Experimental results show that the implementation of the OMP algorithm reported in Ref. 85 achieves +40 $\times$  speed up running on the GTX480 GPU module, controlled by using an Intel(R) Core(TM) i7 CPU.<sup>103,104</sup> GPU-acceleration architecture is a good option to apply single-pixel reconstruction 3D imaging to real-time applications.

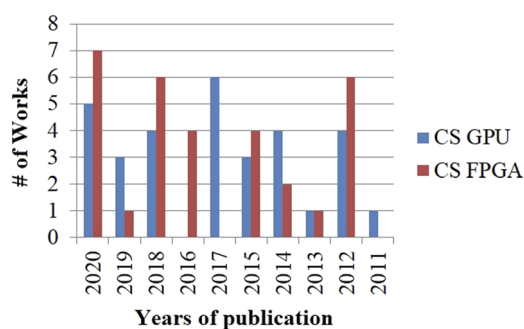
The results of comparing different processing times in applications using GPU platforms, on the one hand, and FPGA modules, on the other, for SPI reconstruction (see Table VI) show that FPGA-based architectures are more efficiently than those based on GPU platforms. During the last few years, the reconstruction of 2D/3D single-pixel images (see Fig. 17) has been increasingly applying the ML approach. The GPU based architectures have been tested in



**FIG. 17.** Number of works reported from 2006 to 2020 with algorithms used for the reconstruction of single-pixel using compressed sensing (CS) techniques with 120 published and TVLA3 with 5 published. The CS<sup>54</sup> methods are more used for single-pixel, which can be adaptive to the other as machine learning (ML)<sup>63</sup> and FSI.<sup>60</sup>

applications based on CNN,<sup>64</sup> while the FPGA architectures have been tested for ML-based approaches, although these do not match the data processing efficiency achieved by GPUs. Due to the fact that it can be easily reconfigured at will, the FPGA accelerator is more flexibly if compared to the GPU and allows for implementation of embedded parallel architectures optimized in terms of clock cycles for the calculation of matrix operations, such as those used for least-squares solution calculations. On the other hand, in GPU-based implementations, kernel parallel operations and shared memory management<sup>85</sup> are possible. Considering the works reported during the last few years, Fig. 18 shows that both approaches have been followed hand to hand, with architecture complementing each other. The efficiency of an SPI system can focus, leaving the spatial resolution and overall quality of the reconstructed single-pixel image aside, on the total time required for signal collection and image reconstruction, especially if developed for video streaming applications, focusing on the elimination of existing bottlenecks in the data transport mechanisms and seeking to accelerate the process of generating SPI-3D images.<sup>101,105</sup>

SPI technique has undergone an impressive evolution in terms of the system architecture it relies on and different reconstruction algorithms used for image reconstruction, ranging from L1-magic<sup>48</sup> proposed in 2007, the orthogonal matching pursuit (OMP)<sup>42</sup> proposed in 2010, the augmented Lagrangian approach and alternating direction algorithm (TVAL3)<sup>42</sup> also proposed in 2010, and NESTA<sup>43,44</sup> proposed in 2011 (see Table II), running on different processor architectures, highlighting the CS that adapts to DMD technologies, of which we saw that it is the most used platform in SPI applications (see Fig. 11). OMP-type algorithms adapt more efficiently in mathematical operations to architectures such as GPUs. Through the applications of sparsity or Nyquist theories,<sup>21</sup> it is possible to define compressed reconstruction strategies, so it is not necessary to make a complete reconstruction of the image using all the projection patterns required for HSI and FSI SPI. Using this fact, it is possible to reduce the reconstruction times simultaneously increasing the efficiency of the acceleration hardware. Based on the data presented in the publications reporting on systems based on both, FPGA and GPU platforms, applied to SPI, it can be foreseen that in the coming years for 3D image applications, it will be necessary

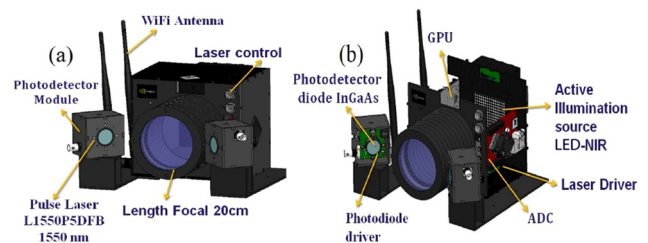


**FIG. 18.** Number of works reported from 2006 to 2020 with algorithms used for the reconstruction of single-pixel CS (using the different type of orthogonal matching algorithm) using hardware acceleration based on FPGA with 31 published and using the paralleled algorithm in GPU with 31 published; both developments in the recent decade gone hand to hand.

to merge both integrated solutions in a single platform to optimize both data collection and reconstruction and processing algorithms. At the level of reconstruction algorithms, the evolution of CS toward interaction with ML can be self-adaptive with the operating conditions of different test scenarios, improving the reconstruction algorithm's efficiency for generation of better quality images and decreasing processing times to reach continuous-time 3D image generation.

### VII. 3D NIR-SPI-BASED MULTI-SPECTRAL CAMERA FOR HARSH ENVIRONMENT APPLICATIONS

An improvement suggested for the 3D SPI-based imaging and ranging system has been the fusion of millimeter-wave RADAR technologies with ToF integrated solutions embedded into SPI photometric stereo technique-oriented systems, enabling unmanned vehicle (UMV) autonomous driving<sup>106</sup> [depicted in Fig. 19(a)]. This hybrid approach allows for continuous-time depth map generation of the surroundings of UMVs with high spatial resolution of objects placed in immediate vicinity or at several meters of distance, and RADAR detected objects placed at much bigger distances, although with much lower spatial resolutions, with an additional advantage offered by the possibility of detecting objects across non-metallic surfaces.<sup>107</sup> Due to the atmosphere's capacity to absorb the wavelengths in the near-infrared spectrum, the best option for generation of single-pixel (or line-sensor based) images for unmanned systems would be to use a structured illumination scheme of front modulation type with an array of LEDs emitting in the NIR part of the spectra at 1550 nm wavelengths [depicted in Fig. 19(b)], which is less sensitive to background noise and undergoes fewer scattering processes when in the presence of particles of different sizes present in fog, rain, snow, or smoke rich environments. The latter will increase the detection range under outdoor conditions or under low-vision conditions (scenarios with dust, fog, rain, or smoke), maximizing the capacity of photometric stereo 3D imaging. A solution is proposed for autonomous driving or flying in unmanned vehicles that should be optimized in terms of dimensions, weight, and power consumption and should yield processing times of below 30 ms to be used for continuous-time decision making. For this, parallel processing based on GPU platforms should be chosen, in combination with OMP-based algorithms<sup>85</sup> or some of its derivations such as



**FIG. 19.** The proposed hyper-spectral camera system proposed, with dimensions of  $11 \times 12 \times 13 \text{ cm}^3$ , 1.3 kg weight, and power consumption of 25 W: (a) the front part of the system is shown, where the photodiode, the focal lens with a focal length of 20 cm, and the Wi-Fi antenna can be observed. (b) The internal design of the hyper-spectral camera is shown where the GPU unit, the analog-to-digital converter (ADC), the active illumination source, the photodiode driver, and the InGaAs SPD can be observed.

TVAL3,<sup>42</sup> or NESTA,<sup>44</sup> to achieve processing time between 30 and 60 frames/s.

### VIII. UAV APPLICATIONS

Commercial success of UAVs has enabled their use in several civilian applications, among which the most notable have been critical infrastructure inspection, aerial video acquisition for cinema, TV and sport events, and precision agriculture.<sup>108</sup> However, UAVs are expected to be deployed under adequate weather conditions. That is, pilots seek to operate under moderate low wind currents, avoiding to fly in cluttered spaces and, in particular, avoiding to fly in cases where the line of sight could be compromised. Examples of the latter could be those of flying in the fog, amid smoke (produced by a fire), or under the rain. While UAV operation heavily relies on a pilot, the UAV may be locally equipped with sensors, enabling it to detect obstacles. These sensors could be based on active pixel acquisition such as stereo RGB cameras, LIDAR, or ToF cameras, including other perception modes not relying on vision, such as radar or auditory perception,<sup>109</sup> which can detect, localize, and classify the audio source just by hearing its audio signal.

Perception is essential not only to aid human piloting of UAVs, but it is also a key component for the development of autonomous UAVs. Hazardous scenarios where humans cannot be physically present are the opportunity to deploy autonomous UAVs. Repetitive and tedious flight operations are another opportunity, for instance, infrastructure inspection, such as warehouse inspection,<sup>110</sup> aerial manipulation,<sup>111</sup> or wind turbine inspection.<sup>112</sup> Furthermore, one of the ultimate goals of autonomous UAV is to achieve agile flight in complex environments<sup>113</sup> and in non-urban environments where GPS cannot be accessed, such as forestry environments.<sup>114</sup>

Among current perception paradigms, active vision-based are perhaps the most popular due to the variety of low-cost, power-efficient, and reduced-size chips commercially available. However, active vision will become ineffective under harsh environments where fog, smoke, or rain may rest visibility of the scene. In recent years, a new paradigm known as event cameras<sup>115</sup> has emerged as an alternative to conventional active vision. Rather than acquiring chromatic information for every pixel in the image grid, an event camera acquires information only in those pixels where a change in intensity has occurred. This enables for the acquisition of a million of “events” per second while also enabling dynamic range. As it has been shown in several works, event cameras could be the way to go in environments with challenging illumination conditions such as with very low light, bright light, or drastic light changes.

It is too early to claim that event cameras will take over conventional active vision as much as it is still an open question whether such sensors will cope with heavy smoke or rain. In contrast, our proposed hyper-spectral system based on SPI aims at extending UAV sensing capabilities by enabling sensing under the harsh environment conditions mentioned before. Figure 20 shows a schematic view of our proposed SPI system mounted on a UAV. Furthermore, we argue that SPI sensing will hugely benefit autonomous UAVs, whose first challenge is that of being capable of sensing its surrounding in order to build a map, which is ultimately used for localization, trajectory planning, and autonomous flight, all of this in addition to sense-and-avoid capabilities during navigation either

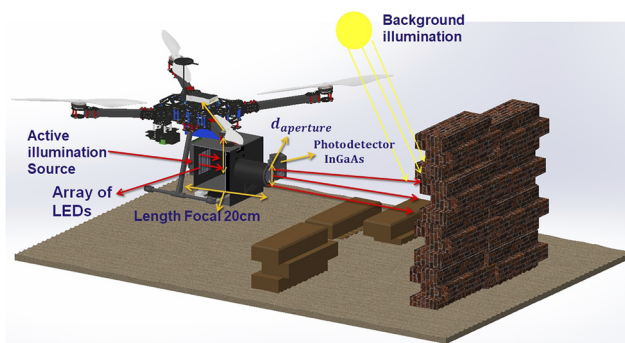


FIG. 20. Schematic view to illustrate the use of our proposed hyper-spectral system carried by a drone and to perform a measurement of energy level. The source of noise from the background lighting, distance, and stages of the optical system are considered.

for static or dynamic objects. Therefore, we argue that SPI systems will contribute to extend usability scenarios of current UAV technology.

### IX. CHALLENGERS OF THE SINGLE-PIXEL TECHNOLOGY IN THE FUTURE

In the last decade, with the increase in the development of technologies as autonomous robots, including self-driving vehicles and UAVs, at the same time, there has been an evolution in techniques to improve vision systems. For that development, different types of sensors such as LIDAR, RADAR, and thermal and IR cameras are used. The system’s vision SPI is a new paradigm to be used as sensor vision that can be adapted to different wavelengths from visible, near-IR, or other long-wavelengths. This advantage allows us

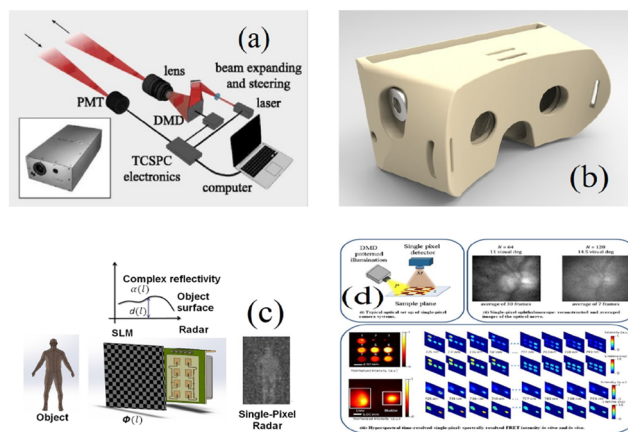


FIG. 21. Overview of new fields for applications of system vision SPI: (a) SPI-LIDAR<sup>29</sup> [Reproduced from Radwell *et al.*, Appl. Phys. Lett. **115**, 231101 (2019) with the permission of AIP Publishing]. (b) VR/AR. (c) SPI-RADAR. (d) SPI-medical<sup>128</sup> [Reprinted with permission from Angelo *et al.*, J. Biomed. Opt. **24**, 071602 (2018). Copyright 2018 Author(s), licensed under a Creative Commons Attribution 4.0 License].

to operate under foggy, rainy, and low-visibility conditions, only changing the source of light without modifying the schematic of configuration SPI. These offer a wide field of applications far from the typical proof of concept that can be seen in the literature, for example, in the development of a robust system of obstacle detection under foggy or rainy condition for application in vehicles<sup>116</sup> [depicted in Fig. 21(a)] and the integration with LIDAR technologies to improve the scene reconstruction quality using an optimized number of samples based on the principle of SPI.<sup>117,118</sup>

SPI applications are not limited to vision sensors, but new applications have also been developed. For example, in VR/AR systems, we can cite the MEMSEye<sup>119</sup> system that works using the SPI and MEMS mirror in the generation of 3D images [depicted in Fig. 21(b)] and applications in holography.<sup>120</sup> Others applications that are studied are adaption as the vision system for ROVs or underwater,<sup>121,122</sup> SPI-RADAR<sup>123</sup> [depicted in Fig. 21(c)], and hyperspectral.<sup>124,125</sup> This later field of application has great importance in the medical<sup>126</sup> area [depicted in Fig. 21(d)], the space exploration,<sup>123</sup> and CubeSat.<sup>127</sup>

## X. CONCLUSIONS

In this work, we have provided a technical review of single-pixel images, initially reviewing the evolution of the first SPI concept's main developments from 1982 to 2020. During this period, it is possible to see an evolution in both the theory behind the sampling and processing approaches, the definition of illumination patterns, and active illumination modulation architectures proposed, conceived for different applications ranging from 2D image generation to the generation of 3D images and depth maps, combining different techniques such as ToF or stereo vision.

Resulting from this review, it is clear that single-pixel cameras and computational GI systems are similar in an optical sense. It should be clarified that there is a distinction between the two of them since single-pixel cameras are often based on structured detection and compressed detection. In contrast, computational GI is often based on structured lighting.

By carrying out a review of the state of the art and focusing on 3D imaging, we can observe that both techniques go hand in hand.

In some applications, such as THz and x-ray detection, the GI is better adapted. However, for applications oriented to unmanned systems, single-pixel has a greater integration capacity, given that SPI enables for the development of low-cost cameras using LED arrays for the generation of 2D/3D images at different wavelengths, even outside the visible spectrum, where focal plane matrix detectors are not available or are prohibitively expensive. A single detector has temporal resolution capabilities superior to the focal plane matrix, which increases the capabilities of the same in IR wavelengths, thus having a greater sensitivity for 3D imaging applications, which can complement other technologies such as radar.

From very early in the development of single-pixel images, ways to reduce the time of the acquired data and the reconstruction of 3D images have been investigated. For that reason, processing techniques have been developed, including orthogonal bases of sampling patterns, compression detection, high-speed spatial light modulation, Machine Learning (ML), wavelet, low-frequency random sampling techniques, and Fourier transforms. ML is a new field of applications that has shown promise in high-speed 3D LIDAR

systems widely used in autonomous vehicles. However, due to LIDAR's high-cost, the single-pixel will be the best solution, with fast processing times and multi-wavelength operating capabilities that can be adapted to rapid detection and classification applications, where it is sufficient to detect signals of characteristic intensity without doing the image reconstruction.

Based on data found in the works we reviewed, the single-pixel algorithms are leveraged by parallel implementations on FPGA and GPU, thus enabling 3D applications near real-time (the operation frequency of 30 Hz or more). Furthermore, we can see an evolution of the conventional algorithms toward the incorporation of techniques into the artificial intelligence field such as machine learning.

Finally, single-pixel offers the possibility to create a new unique low-cost sensing technology, expanding a new field of research and applications, in particular, applications requiring low-cost sensors to be used in outdoor environments with harsh conditions and to be carried by unmanned systems, for instance, UAVs having to fly under the rain, fog, or amid smoke produced. Currently, UAVs struggle to navigate under such environment conditions; however, SPI could enable a safe navigation by enabling UAVs to sense even under such adverse conditions. If this is the case, it opens the door for a plethora of applications.

It is important to keep in mind that SPI is not only limited to be used as a vision sensor for vehicles or UAVs, but there are other fields of applications such as medicine and other applications that are undergoing its integration with RADAR, LIDAR, the field of space exploration, VR/AR, and ROV system. Therefore, we foresee that SPI will become a common vision sensor due to its capacity of integration with other technologies.

## ACKNOWLEDGMENTS

C.A.O.Q. acknowledges Consejo Nacional de Ciencia y Tecnología (CONACYT) for the scholarship under Grant No. CVU: 661331.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

All the authors contributed equally to this work.

## DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## REFERENCES

- <sup>1</sup>D. Durini and D. Arutinov, "2-operational principles of silicon image sensors," in *High Performance Silicon Imaging*, Woodhead Publishing Series in Electronic and Optical Materials, 2nd ed., edited by D. Durini (Woodhead Publishing, 2020), pp. 25–73.
- <sup>2</sup>N. Ben-Yosef and G. Sirat, "Real-time spatial filtering utilizing the piezoelectric-elasto-optic effect," *Opt. Acta: Int. J. Opt.* **29**, 419–423 (1982).

- <sup>3</sup>D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
- <sup>4</sup>D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," *Proc. SPIE* **6065**, 606509 (2006).
- <sup>5</sup>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**, 83–91 (2008).
- <sup>6</sup>S. S. Welsh, M. P. Edgar, R. Bowman, P. Jonathan, B. Sun, and M. J. Padgett, "Fast full-color computational imaging with single-pixel detectors," *Opt. Express* **21**, 23068–23074 (2013).
- <sup>7</sup>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**, 844–847 (2013).
- <sup>8</sup>P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, "Dual photography," *ACM Trans. Graphics* **24**, 745–755 (2005).
- <sup>9</sup>E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inf. Theory* **52**, 5406–5425 (2006).
- <sup>10</sup>E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Probl.* **23**, 969–985 (2007).
- <sup>11</sup>G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, "Photon counting compressive depth mapping," *Opt. Express* **21**, 23822–23837 (2013).
- <sup>12</sup>G. A. Howland, P. B. Dixon, and J. C. Howell, "Photon-counting compressive sensing laser radar for 3D imaging," *Appl. Opt.* **50**, 5917–5920 (2011).
- <sup>13</sup>D. Shrekenhamer, C. M. Watts, and W. J. Padilla, "Terahertz single pixel imaging with an optically controlled dynamic spatial light modulator," *Opt. Express* **21**, 12507–12518 (2013).
- <sup>14</sup>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**, 5834 (2014).
- <sup>15</sup>H. Dai, G. Gu, W. He, F. Liao, J. Zhuang, X. Liu, and Q. Chen, "Adaptive compressed sampling based on extended wavelet trees," *Appl. Opt.* **53**, 6619–6628 (2014).
- <sup>16</sup>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).
- <sup>17</sup>S. M. Hornett, R. I. Stantchev, M. Z. Vardaki, C. Beckerleg, and E. Hendry, "Subwavelength terahertz imaging of graphene photoconductivity," *Nano Lett.* **16**, 7019–7024 (2016).
- <sup>18</sup>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**, 2998–3005 (2017).
- <sup>19</sup>C. Higham, R. Murray-Smith, M. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Sci. Rep.* **8**, 2369 (2018).
- <sup>20</sup>K. Nitta, Y. Yano, C. Kitada, and O. Matoba, "Fast computational ghost imaging with laser array modulation," *Appl. Sci.* **9**, 4807 (2019).
- <sup>21</sup>W.-K. Yu, "Super sub-Nyquist single-pixel imaging by means of cake-cutting Hadamard basis sort," *Sensors* **19**, 4122 (2019).
- <sup>22</sup>Q. Yi, L. Z. Heng, L. Liang, Z. Guangcan, C. F. Siong, and Z. Guangya, "Hadamard transform-based hyperspectral imaging using a single-pixel detector," *Opt. Express* **28**, 16126–16139 (2020).
- <sup>23</sup>A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato, "Ghost imaging with thermal light: Comparing entanglement and classical correlation," *Phys. Rev. Lett.* **93**, 093602 (2004).
- <sup>24</sup>A. Valencia, G. Scarcelli, M. D'Angelo, and Y. Shih, "Two-photon imaging with thermal light," *Phys. Rev. Lett.* **94**, 063601 (2005).
- <sup>25</sup>J. H. Shapiro, "Computational ghost imaging," *Phys. Rev. A* **78**, 061802 (2008).
- <sup>26</sup>O. Katz, Y. Bromberg, and Y. Silberberg, "Compressive ghost imaging," *Appl. Phys. Lett.* **95**, 131110 (2009).
- <sup>27</sup>Y. Bromberg, O. Katz, and Y. Silberberg, "Ghost imaging with a single detector," *Phys. Rev. A* **79**, 053840 (2009).
- <sup>28</sup>F. Ferri, D. Magatti, L. Lugiato, and A. Gatti, "Differential ghost imaging," *Phys. Rev. Lett.* **104**, 253603 (2010).
- <sup>29</sup>N. Radwell, S. Johnson, M. Edgar, C. Higham, R. Murray-Smith, and M. Padgett, "Deep learning optimized single-pixel LiDAR," *Appl. Phys. Lett.* **115**, 231101 (2019).
- <sup>30</sup>Z. Zhang, X. Ma, and J. Zhong, "Single-pixel imaging by means of Fourier spectrum acquisition," *Nat. Commun.* **6**, 6225 (2015).
- <sup>31</sup>J.-Z. Yang, M.-F. Li, X.-X. Chen, W.-K. Yu, and A.-N. Zhang, "Single-photon quantum imaging via single-photon illumination," *Appl. Phys. Lett.* **117**, 214001 (2020).
- <sup>32</sup>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**, 2427–2434 (2018).
- <sup>33</sup>Z. Zhang and J. Zhong, "Three-dimensional single-pixel imaging with far fewer measurements than effective image pixels," *Opt. Lett.* **41**, 2497–2500 (2016).
- <sup>34</sup>D. Stork and P. Gill, "Lensless ultra-miniature CMOS computational imagers and sensors," **2013**, 186–190.
- <sup>35</sup>V. Boominathan, J. K. Adams, M. S. Asif, B. W. Avants, J. T. Robinson, R. G. Baraniuk, A. C. Sankaranarayanan, and A. Veeraraghavan, "Lensless imaging: A computational renaissance," *IEEE Signal Process. Mag.* **33**, 23–35 (2016).
- <sup>36</sup>L. Chen, Z. Zhou, L. Deng, N. Xi, B. Song, Y. Yang, Y. Hao, and Z. Sun, "Single image super resolution infrared camera using carbon nanotube photodetector," in *2015 IEEE Sensors* (IEEE, 2015), pp. 1–4.
- <sup>37</sup>E. Salvador-Balaguer, P. Latorre-Carmona, C. Chabert, F. Pla, J. Lancis, and E. Tajahuerce, "Low-cost single-pixel 3D imaging by using an LED array," *Opt. Express* **26**, 15623–15631 (2018).
- <sup>38</sup>G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: A review," *Opt. Express* **28**, 28190–28208 (2020).
- <sup>39</sup>M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," *Nat. Photonics* **13**, 13–20 (2019).
- <sup>40</sup>L. Bian, J. Suo, Q. Dai, and F. Chen, "Experimental comparison of single-pixel imaging algorithms," *J. Opt. Soc. Am. A* **35**, 78–87 (2018).
- <sup>41</sup>M.-J. Sun and J.-M. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: A brief review," *Sensors* **19**, 732 (2019).
- <sup>42</sup>L. Chengbo, "An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing," M.S. thesis, Rice University, 2010.
- <sup>43</sup>S. Becker, J. Bobin, and E. J. Candès, "Nesta: A fast and accurate first-order method for sparse recovery," *SIAM J. Imaging Sci.* **4**, 1–39 (2011).
- <sup>44</sup>P. G. Vaz, D. Amaral, L. F. R. Ferreira, M. Morgado, and J. Cardoso, "Image quality of compressive single-pixel imaging using different Hadamard orderings," *Opt. Express* **28**, 11666–11681 (2020).
- <sup>45</sup>M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.* **55**, 5695–5702 (2007).
- <sup>46</sup>H.-C. Liu, "Imaging reconstruction comparison of different ghost imaging algorithms," *Sci. Rep.* **10**, 14626 (2020).
- <sup>47</sup>J. Cao, D. Zhou, F. Zhang, H. Cui, Y. Zhang, and Q. Hao, "A novel approach of parallel retina-like computational ghost imaging," *Sensors* **20**, 7093 (2020).
- <sup>48</sup>E. Hale, W. Yin, and Y. Zhang, "A fixed-point continuation method for  $l_1$ -regularized minimization with applications to compressed sensing," CAAM Technical Report No. TR07-07, 2007.
- <sup>49</sup>B. Sun, S. S. Welsh, M. P. Edgar, J. H. Shapiro, and M. J. Padgett, "Normalized ghost imaging," *Opt. Express* **20**, 16892–16901 (2012).
- <sup>50</sup>K.-H. Luo, B.-Q. Huang, W.-M. Zheng, and L.-A. Wu, "Nonlocal imaging by conditional averaging of random reference measurements," *Chin. Phys. Lett.* **29**, 074216 (2012).
- <sup>51</sup>W. Gong, "High-resolution pseudo-inverse ghost imaging," *Photonics Res.* **3**, 234–237 (2015).
- <sup>52</sup>S. M. M. Khamoushi, Y. Nosrati, and S. H. Tavassoli, "Sinusoidal ghost imaging," *Opt. Lett.* **40**, 3452–3455 (2015).
- <sup>53</sup>L. Wang and S. Zhao, "Fast reconstructed and high-quality ghost imaging with fast Walsh–Hadamard transform," *Photonics Res.* **4**, 240–244 (2016).
- <sup>54</sup>M.-J. Sun, M. Tong, M. Edgar, M. Padgett, and N. Radwell, "A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging," *Sci. Rep.* **7**, 3464 (2017).

- <sup>55</sup>W.-K. Yu and Y.-M. Liu, "Single-pixel imaging with origami pattern construction," *Sensors* **19**, 5135 (2019).
- <sup>56</sup>Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Hadamard single-pixel imaging versus Fourier single-pixel imaging," *Opt. Express* **25**, 19619–19639 (2017).
- <sup>57</sup>Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Fast Fourier single-pixel imaging via binary illumination," *Sci. Rep.* **7**, 12029 (2017).
- <sup>58</sup>M. Xi, H. Chen, Y. Yuan, G. Wang, Y. He, Y. Liang, J. Liu, H. Zheng, and Z. Xu, "Bi-frequency 3D ghost imaging with Haar wavelet transform," *Opt. Express* **27**, 32349–32359 (2019).
- <sup>59</sup>S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 674–693 (1989).
- <sup>60</sup>G. Strang, "Wavelet transforms versus Fourier transforms," *Bull. Am. Math. Soc.* **28**, 288–305 (1993).
- <sup>61</sup>G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge Press, 1996).
- <sup>62</sup>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**, 13269–13278 (2020).
- <sup>63</sup>J. Zhang, C. Zhao, and W. Gao, "Optimization-inspired compact deep compressive sensing," *IEEE J. Sel. Top. Signal Process.* **14**, 765–774 (2020).
- <sup>64</sup>W. Shi, F. Jiang, S. Zhang, and D. Zhao, "Deep networks for compressed image sensing," *2017 IEEE International Conference on Multimedia and Expo (ICME)* (2017), pp. 877–882.
- <sup>65</sup>Y. Wang, J. Suo, J. Fan, and Q. Dai, "Hyperspectral computational ghost imaging via temporal multiplexing," *IEEE Photonics Technol. Lett.* **28**, 288–291 (2016).
- <sup>66</sup>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**, 113901 (2016).
- <sup>67</sup>S. Augustin, S. Frohmann, P. Jung, and H.-W. Hübers, "Mask responses for single-pixel terahertz imaging," *Sci. Rep.* **8**, 4886 (2018).
- <sup>68</sup>L. Zanotto, R. Piccoli, J. Dong, D. Caraffini, R. Morandotti, and L. Razzari, "Time-domain terahertz compressive imaging," *Opt. Express* **28**, 3795–3802 (2020).
- <sup>69</sup>M. Wenwen, S. Dongfeng, H. Jian, Y. Kee, W. Yingjian, and F. Chengyu, "Sparse Fourier single-pixel imaging," *Opt. Express* **27**, 31490–31503 (2019).
- <sup>70</sup>Lc-slm, at URL: <http://www.lasercomponents.com/us/product/lc-2012-lcd-slm/>; accessed March 2021.
- <sup>71</sup>Leds array, at URL: <http://www.teleopto.com/teleopto/led-array-system/>; accessed March 2021.
- <sup>72</sup>K. Pieper, A. Bergmann, R. Dengler, and C. Rockstuhl, "Using a pseudo-thermal light source to teach spatial coherence," *Eur. J. Phys.* **39**, 045303 (2018).
- <sup>73</sup>Y.-X. Li, W.-K. Yu, J. Leng, and S.-F. Wang, "Pseudo-thermal imaging by using sequential-deviations for real-time image reconstruction," *Opt. Express* **27**, 35166–35181 (2019).
- <sup>74</sup>Dmd, at URL: <http://www.ti.com/dlp-chip/getting-started.html>; accessed March 2021.
- <sup>75</sup>S. Ahmadi Kandjani, R. Kheradmand, and N. Dadashzadeh, "Ghost imaging with pseudo-thermal light," in *2011 13th International Conference on Transparent Optical Networks* (IEEE, 2011), pp. 1–4.
- <sup>76</sup>Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3D single-pixel video," *J. Opt.* **18**, 035203 (2016).
- <sup>77</sup>M. Wang, M.-J. Sun, and C. Huang, "Single-pixel 3D reconstruction via a high-speed LED array," *J. Phys.: Photonics* **2**, 025006 (2020).
- <sup>78</sup>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. Express* **27**, 3731–3745 (2019).
- <sup>79</sup>M. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Natl. Bur. Stand.* **49**, 409–435 (1952).
- <sup>80</sup>K. Guo, S. Jiang, and G. Zheng, "Multilayer fluorescence imaging on a single-pixel detector," *Biomed. Opt. Express* **7**, 2425–2431 (2016).
- <sup>81</sup>L. Bian, J. Suo, J. Chung, X. Ou, C. Yang, F. Chen, and Q. Dai, "Fourier ptychographic reconstruction using Poisson maximum likelihood and truncated Wirtinger gradient," *Sci. Rep.* **6**, 27384 (2016).
- <sup>82</sup>H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike*, edited by E. Parzen, K. Tanabe, and G. Kitagawa (Springer New York, New York, 1998), pp. 199–213.
- <sup>83</sup>A. Yang, Z. Zhou, A. Balasubramanian, S. Sastry, and Y. Ma, "Fast 11-minimization algorithms for robust face recognition," *IEEE Transactions on Image Processing* (IEEE, 2013), Vol. 22.
- <sup>84</sup>H. Zhu, W. Chen, and Y. Wu, "Efficient implementations for orthogonal matching pursuit," *Electronics* **9**, 1507 (2020).
- <sup>85</sup>C. O. Quero, D. Durini, R. Ramos-Garcia, J. Rangel-Magdaleno, and J. Martinez-Carranza, "Hardware parallel architecture proposed to accelerate the orthogonal matching pursuit compressive sensing reconstruction," *Proc. SPIE* **11396**, 113960N (2020).
- <sup>86</sup>S.-H. Hsieh, C.-S. Lu, and S.-C. Pei, "Fast OMP: Reformulating OMP via iteratively refining  $\ell_2$ -norm solutions," in *2012 IEEE Statistical Signal Processing Workshop (SSP)* (IEEE, 2012), pp. 189–192.
- <sup>87</sup>A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, "Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting," *Appl. Opt.* **48**, 6241–6251 (2009).
- <sup>88</sup>A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat. Commun.* **3**, 745 (2012).
- <sup>89</sup>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.
- <sup>90</sup>R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.* **19**, 191139 (1980).
- <sup>91</sup>A. McCarthy, N. J. Krichel, N. R. Gemmill, X. Ren, M. G. Tanner, S. N. Dorenbos, V. Zwiller, R. H. Hadfield, and G. S. Buller, "Kilometre-range, high resolution depth imaging using 1560 nm wavelength single-photon detection," in *2013 Conference on Lasers and Electro-Optics–International Quantum Electronics Conference* (Optical Society of America, 2013).
- <sup>92</sup>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**, 12010 (2016).
- <sup>93</sup>L. Gao, J. Liang, C. Li, and L. V. Wang, "Single-shot compressed ultrafast photography at one hundred billion frames per second," *Nature* **516**, 74 (2014).
- <sup>94</sup>L. Jin, R. Aimin, and Z. Jingnan, "3D reconstruction by perspective shape from shading using linearized triangular element surface model," in *2006 International Conference on Mechatronics and Automation* (IEEE, 2006), pp. 1763–1768.
- <sup>95</sup>Laser safety facts, at URL: <https://www.lasersafetyfacts.com/laserclasses.html>; accessed March 2021.
- <sup>96</sup>P. Caramazza, A. Boccolini, D. Buschek, M. Hullin, C. F. Higham, R. Henderson, R. Murray-Smith, and D. Faccio, "Neural network identification of people hidden from view with a single-pixel, single-photon detector," *Sci. Rep.* **8**, 11945 (2018).
- <sup>97</sup>R. Pandharkar, A. Velten, A. Bardagjy, E. Lawson, M. Bawendi, and R. Raskar, "Estimating motion and size of moving non-line-of-sight objects in cluttered environments," in *CVPR 2011* (IEEE, 2011), pp. 265–272.
- <sup>98</sup>R. Warburton, S. Chan, G. Garipey, Y. Altmann, S. McLaughlin, J. Leach, and D. Faccio, "Real-time tracking of hidden objects with single-pixel detectors," in *Imaging and Applied Optics* (Optical Society of America, 2016), p. IT4E.2.
- <sup>99</sup>A. Yuniarti and N. Suciati, "A review of deep learning techniques for 3D reconstruction of 2D images," in *2019 12th International Conference on Information Communication Technology and System (ICTS)* (IEEE, 2019), pp. 327–331.
- <sup>100</sup>I. Yasri, "An FPGA based hardware accelerator for real time video segmentation system," in *2011 International Conference on Advanced Computer Science and Information Systems* (IEEE, 2011), pp. 63–68.
- <sup>101</sup>S. Bi, N. Xi, K. W. C. Lai, and X. Pan, "Design and implementation for image reconstruction of compressivesensing using fpga," in *2013 IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems* (IEEE, 2013), pp. 320–325.

- <sup>102</sup>A. Mase, Y. Quan, Y. Li, X. Gao, and M. Xing, "FPGA implementation of real-time compressive sensing with partial Fourier dictionary," *Int. J. Antennas Propag.* **2016**, 1671687.
- <sup>103</sup>A. Kulkarni and T. Mohsenin, "Accelerating compressive sensing reconstruction OMP algorithm with CPU, GPU, FPGA and domain specific many-core," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2015), pp. 970–973.
- <sup>104</sup>Y. Fang, L. Chen, J. Wu, and B. Huang, "GPU implementation of orthogonal matching pursuit for compressive sensing," in *2011 IEEE 17th International Conference on Parallel and Distributed Systems* (IEEE, 2011), pp. 1044–1047.
- <sup>105</sup>C. A. Osorio Quero, D. D. Romero, R. Ramos-Garcia, J. de Jesus Rangel-Magdaleno, and J. Martinez-Carranza, "Towards a 3D vision system based on single-pixel imaging and indirect time-of-flight for drone applications," in *2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)* (IEEE, 2020), pp. 1–6.
- <sup>106</sup>C. O. Quero, D. Durini, R. Ramos-Garcia, J. Rangel-Magdaleno, and J. Martinez-Carranza, "Evaluation of a 3D imaging vision system based on a single-pixel InGaAs detector and the time-of-flight principle for drones," *Proc. SPIE* **11402**, 114020T (2020).
- <sup>107</sup>G. Gennarelli, F. Soldovieri, and M. Amin, "Radar for indoor surveillance: State of art and perspectives," *Proc. SPIE* **11059**, 1105903 (2019).
- <sup>108</sup>S. Robotics, Robotics 2020 multi-annual roadmap for robotics in Europe, SPARC Robotics, EU-Robotics AISBL, The Hague, The Netherlands; accessed February 5, 2018, 2016.
- <sup>109</sup>J. Martinez-Carranza and C. Rascon, "A review on auditory perception for unmanned aerial vehicles," *Sensors* **20**, 7276 (2020).
- <sup>110</sup>J. Martinez-Carranza, R. Bostock, S. Willcox, I. Cowling, and W. Mayol-Cuevas, "Indoor MAV auto-retrieval using fast 6D relocalisation," *Adv. Rob.* **30**, 119–130 (2016).
- <sup>111</sup>A. Lopez Luna, I. Cruz Vega, and J. Martinez-Carranza, "Vertical surface contact with a micro air vehicle," *Int. J. Micro Air Veh.* **12**, (2020).
- <sup>112</sup>R. Parlange, J. Martinez-Carranza, L. Sucar, B. Ren, and S. Watkins, "Vision-based autonomous navigation for wind turbine inspection using an unmanned aerial vehicle," in *10th International Micro Air Vehicle Conference and Competition* (IMAV, 2018), pp. 283–288.
- <sup>113</sup>H. Moon, J. Martinez-Carranza, T. Cieslewski, M. Faessler, D. Falanga, A. Simovic, D. Scaramuzza, S. Li, M. Ozo, C. De Wagter *et al.*, "Challenges and implemented technologies used in autonomous drone racing," *Intell. Serv. Rob.* **12**, 137–148 (2019).
- <sup>114</sup>S. Dionisio-Ortega, L. O. Rojas-Perez, J. Martinez-Carranza, and I. Cruz-Vega, "A deep learning approach towards autonomous flight in forest environments," in *2018 International Conference on Electronics, Communications and Computers (CONIELECOMP)* (IEEE, 2018), pp. 139–144.
- <sup>115</sup>G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: A survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence* (IEEE, 2020).
- <sup>116</sup>R. Tobin, A. Halimi, A. McCarthy, P. J. Soan, and G. S. Buller, "Robust real-time 3D imaging of moving scenes through atmospheric obscurant using single-photon LiDAR," *Sci. Rep.* **11**, 11236 (2021).
- <sup>117</sup>S. D. Johnson, N. Radwell, M. P. Edgar, C. Higham, R. Murray-Smith, and M. J. Padgett, "Single-pixel lidar with deep learning optimised sampling," in *Conference on Lasers and Electro-Optics* (Optical Society of America, 2020), p. ATu3T.5.
- <sup>118</sup>F. Magalhães, M. V. Correia, F. Farahi, J. Pereira do Carmo, and F. M. Araújo, "Cyclops: Single-pixel imaging lidar system based on compressive sensing," *Proc. SPIE* **10563**, 105633U (2017).
- <sup>119</sup>V. Milanović, A. Kasturi, J. Yang, and F. Hu, "A fast single-pixel laser imager for VR/AR headset tracking," *Proc. SPIE* **10116**, 101160E (2017).
- <sup>120</sup>L. Martínez-León, P. Clemente, Y. Mori, V. Climent, J. Lancis, and E. Tajahuerce, "Single-pixel digital holography with phase-encoded illumination," *Opt. Express* **25**, 4975–4984 (2017).
- <sup>121</sup>M. Li, A. Mathai, S. L. H. Lau, J. W. Yam, X. Xu, and X. Wang, "Underwater object detection and reconstruction based on active single-pixel imaging and super-resolution convolutional neural network," *Sensors* **21**, 313 (2021).
- <sup>122</sup>M. Le, G. Wang, H. Zheng, J. Liu, Y. Zhou, and Z. Xu, "Underwater computational ghost imaging," *Opt. Express* **25**, 22859–22868 (2017).
- <sup>123</sup>B. Fürsich, R. Bamler, S. Augustin, H.-W. Hübers, and X. X. Zhu, "Towards single-pixel FMCW radar reconstruction," in *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)* (IEEE, 2016), pp. 95–99.
- <sup>124</sup>F. Magalhaes, F. M. Araújo, M. Correia, M. Abolbashari, and F. Farahi, "High-resolution hyperspectral single-pixel imaging system based on compressive sensing," *Opt. Eng.* **51**, 071406 (2012).
- <sup>125</sup>C. Tao, H. Zhu, X. Wang, S. Zheng, Q. Xie, C. Wang, R. Wu, and Z. Zheng, "Compressive single-pixel hyperspectral imaging using RGB sensors," *Opt. Express* **29**, 11207–11220 (2021).
- <sup>126</sup>S. Jin, W. Hui, Y. Wang, K. Huang, Q. Shi, C. Ying, D. Liu, Q. Ye, W. Zhou, and J. Tian, "Hyperspectral imaging using the single-pixel Fourier transform technique," *Sci. Rep.* **7**, 45209 (2017).
- <sup>127</sup>C. Lastrì, V. Guzzi Nardino, L. Palombi, and V. Raimondi, "Exploring the potential of compressive sensing and super-resolution for space applications in the MIR-TIR," *Proceedings* **27**, 35 (2019).
- <sup>128</sup>J. Angelo, S.-J. Chen, M. Ochoa Mendoza, U. Sunar, S. Gioux, and X. Intes, "Review of structured light in diffuse optical imaging," *J. Biomed. Opt.* **24**, 071602 (2018).