Fast hyperspectral single-pixel imaging

Guilherme Beneti Martins, Laurent Mahieu-Williame, Antonio Mur, Bruno Montcel, Nicolas Ducros

Fast hyperspectral single-pixel imaging

GUILHERME BENETI MARTINS [1], LAURENT MAHIEU-WILLIAME,[1] ANTONIO LORENTE MUR[1], BRUNO MONTCEL[1], AND NICOLAS DUCROS[1]

[1] *Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR 5220, U1294, F-69621, LYON, France*

[*] *nicolas.ducros@creatis.insa-lyon.fr*

Abstract: Hyperspectral imaging is a major tool in modern science, which relies on a compromise between spatial resolution, spectral resolution, and imaging speed. Inspired by single-pixel imaging, we propose a versatile system that enables the fast acquisition of high-spectral-resolution hypercubes. Our computational hyperspectral imaging device is composed of a compact fiber spectrometer and a digital micromirror device (DMD). By uploading a set of Hadamard patterns onto the DMD, our system acquires $64 \times 64 \times 2048$ pixel hypercubes with a spectral resolution of 2.3 nm in less that 2 s. We show that this time can be further reduced by reconstructing hypercubes from accelerated acquisitions that exploit only a few DMD patterns. In particular, we demonstrate that a deep expectation maximization network (EM-Net) can solve this inverse problem for several acceleration factors. 8-fold acceleration enables the achievement of reconstructions with moderate spatial degradation for low frequency images. Our system allows a high degree of flexibility in the choice of spatial resolution and imaging speed, which can be easily adapted to the target application. To foster research in this field, we have made our image reconstruction algorithms, acquisition software, and several raw datasets publicly available.

1. Introduction

Spectral imaging is a major tool of modern science, with applications in astronomy, environmental monitoring, food processing, agriculture, and biomedical imaging. Approaches for spectral imaging are usually categorized as pushbroom, filter-based, or snapshot [1,2]. The pushbroom and filter-based methods are scanning techniques that require multiple measurements to acquire a full (x, y, λ) hypercube. Pushbroom methods acquire one (y, λ) slice at a time and require scanning along the x-axis [3]. Filter-based setups acquire an (x, y) image for one spectral band, with the full hypercube obtained from a sequence of measurements by rotating a filter wheel or monitoring electronically tunable filters [4]. Both pushbroom and filter-based approaches suffer from low throughput as only a small region of the hypercube is measured at a time. Moreover, the spatial (e.g., for pushbroom) or spectral (e.g., for filters) resolutions are linear in proportion to the number of measurements, and hence they are either slow or low resolution. These limitations have led to snapshot methods, which we categorize into hardware and computational methods.

Hardware snapshot methods are remapping strategies where the three-dimensional (3D) hypercube is mapped onto a 2D sensor. Different mapping strategies have led to variants, which include the use of mirror arrays, fiber bundles, and lenslet arrays (see review [5]). Hardware snapshot methods are widespread in astronomy and remote sensing [6] and have recently been introduced in biomedical imaging [7]. While the first generations of hyperspectral imagers were hardware-driven, the most recent snapshot imagers are computational, i.e., they rely on algorithms that reconstruct a hypercube from a few raw measurements. The field of computational imaging has been very active over the past 10 years [8]. Coded aperture snapshot spectral imagers and their different variants exploit a diffractive element with a programmable mask, such that each raw measurement gives access to an oblique projection of the hypercube [9]. Miniature ultra-spectral

imaging uses a liquid crystal phase retarder to multiplex the spectral domain [10]. The spectral DiffuserCam is a multispectral filter array where the optics are replaced by a diffuser, such that each pixel on the sensor can 'see' the whole field of view [11]. The coupling of a compressive spectrometer with a confocal microscope allows for high-sensitivity Raman imaging [12].

However, snapshot imagers suffer from an inherent trade-off between the spatial and the spectral dimensions. Computational snapshot imagers also require reconstruction algorithms that may have long computation times (e.g., several tens of seconds or minutes depending on the dimension of the hypercube).

Hyperspectral single-pixel imaging is a generalization of single-pixel imaging that allows high-spectral-resolution hypercubes to be obtained [13, 14]. From a set of spectra obtained using a set of spatial light patterns, the hypercube is recovered by a reconstruction algorithm [15, 16]. Compared to snapshot imagers, this approach requires more measurements; however, it leads to an excellent spectral resolution, as the spectral dimension is not multiplexed. Recently, fast single-pixel imaging has been demonstrated at several tens of frames in the visible range [17] and terahertz range [18]. Single-pixel imaging also benefits from deep reconstruction methods that enable fast reconstruction while outperforming handcrafted prior-based methods [19–21]. However, these systems have no spectral resolution.

Here, we propose and evaluate a fast hyperspectral imaging device, which relies on a digital micromirror device (DMD) and a compact spectrometer. The DMD displays a sequence of patterns taken from a Hadamard basis and a spectrum is acquired for each pattern. The hypercube is recovered from the set of raw spectra. Our system typically operates at up to 0.1 Hz for 2,048 channels. However, accelerated acquisitions, that display a reduced number of patterns on the DMD, can image at up to 0.7 Hz. To the best of our knowledge, this represents an unprecedented speed for this spectral resolution. Following the recent trend of physics-informed deep learning [22], we consider an explainable reconstruction method based on deep learning. Our deep reconstruction allows for fast reconstructions (e.g., hundreds of milliseconds). Upon acceptance, our reconstruction method will be made available in the Python toolbox SPyRiT [23]. We will also make our experimental datasets and acquisition software publicly available.

2. Methods

The proposed computational framework is depicted in Fig. 1. To acquire a 3D hypercube using a 2D sensor, we acquire multiple pixels at the same time by shaping the light with a DMD. After a sequence of spectra has been taken by using different DMD patterns, we feed the raw spectra into a deep reconstruction algorithm that recovers the hypercube.

2.1. Image formation model

Let $\hat{\boldsymbol{M}} \in \mathbb{R}^{2K\times\Lambda}$ represent the raw measurements, where $2K$ is the number of DMD patterns and Λ the number of spectral channels provided by the spectrometer. Let $\boldsymbol{P} \in \mathbb{R}^{2K\times N}$ be the matrix that contains the DMD patterns, where N is the number of (spatial) pixels in each pattern and $\boldsymbol{F} \in \mathbb{R}^{N\times\Lambda}$ represents the 3D hypercube. We model the acquisition process as linear measurements corrupted by Poissonian-Gaussian noise [24]

$$\hat{\boldsymbol{M}} \sim g\,\mathcal{P}(\boldsymbol{PF}) + \mathcal{N}(\mu_{\text{dark}}, \sigma^2_{\text{dark}}) \tag{1}$$

where $\mathcal{P}$ and $\mathcal{N}$ are the Poisson and Gaussian distributions, g represents the system gain (in counts/electron), μ_{dark} is the dark current (in counts), and σ_{dark} is the dark noise (in counts).

The light patterns are taken from a Hadamard basis and split into positive and negative parts to be uploaded onto the DMD [25]. In notations, we have $\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_+ \\ \boldsymbol{P}_- \end{bmatrix}$ where $\boldsymbol{P}_+ \in \mathbb{R}_+^{K\times N}$ and $\boldsymbol{P}_- \in \mathbb{R}_+^{K\times N}$ are the positive and negative part of Hadamard patterns, i.e., $\boldsymbol{P}_+ - \boldsymbol{P}_- = \boldsymbol{SH}$, where

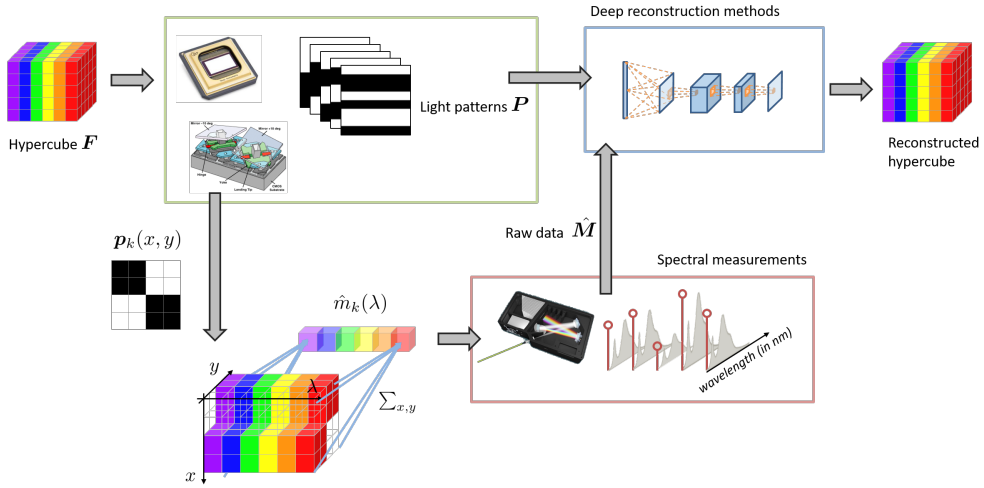

Fig. 1. Hyperspectral single-pixel imaging principle. The hypercube $\boldsymbol{F} \in \mathbb{R}^{N\times\Lambda}$ is sent to a compact spectrophotometer via a digital micromirror device (DMD). A sequence of $2K$ light patterns $\boldsymbol{P} \in \mathbb{R}^{2K\times N}$ is uploaded onto the DMD, leading to the measurement of the $2K$ raw spectra $\hat{\boldsymbol{M}} \in \mathbb{R}^{2K\times\Lambda}$. A deep reconstruction method is then used to reconstruct the hypercube in the case $K < N$.

$\boldsymbol{H} \in \mathbb{R}^{N\times N}$ is the Walsh-Hadamard basis and $\boldsymbol{S} \in \{0,1\}^{K\times N}$ is a subsampling matrix that retains some of the rows of $\boldsymbol{H}$. In the following, we denote the retained Hadamard patterns by $\boldsymbol{H}_\downarrow = \boldsymbol{S}\boldsymbol{H}$.

We finally preprocess the raw measurements $\hat{\boldsymbol{M}} = \begin{bmatrix} \hat{\boldsymbol{M}}_+ \\ \hat{\boldsymbol{M}}_- \end{bmatrix}$ to compensate for splitting

$$\boldsymbol{M} = \hat{\boldsymbol{M}}_+ - \hat{\boldsymbol{M}}_- \tag{2}$$

where $\boldsymbol{M}_+ \in \mathbb{R}^{K\times N}$ and $\boldsymbol{M}_- \in \mathbb{R}^{K\times N}$ correspond to the measurements obtained with the positive and negative patterns, respectively. Therefore, the preprocessed measurements $\boldsymbol{M}$ are Hadamard coefficients, in the sense that $\mathbb{E}(\boldsymbol{M}) = \boldsymbol{H}_\downarrow \boldsymbol{F}$, where $\mathbb{E}$ denotes the expectation. Note that the problem is separable across the spectral dimension, i.e., $\mathbb{E}(\boldsymbol{m}_\lambda) = \boldsymbol{H}_\downarrow \boldsymbol{f}_\lambda$, $1 \le \lambda \le \Lambda$, where $\boldsymbol{m}_\lambda \in \mathbb{R}^K$ and $\boldsymbol{f}_\lambda \in \mathbb{R}^N$ are the λ-th column of $\boldsymbol{M}$ and $\boldsymbol{P}$, respectively. Therefore, the spectral resolution of the hypercube is given directly by the spectral resolution of the spectrometer, while its spatial resolution depends only on the light patterns and our ability to recover $\boldsymbol{f}_\lambda$ from $\boldsymbol{m}_\lambda$.

2.2. Image reconstruction

In the case $K = N$, the hypercube can be reconstructed in the least squares sense as

$$\boldsymbol{F} = \frac{1}{N}\boldsymbol{H}^\top \boldsymbol{M} \tag{3}$$

However, when only a few patterns are considered to limit the acquisition time, i.e., when $K < N$, we propose to reconstruct each λ-slice of the hypercube independently by computing the maximum a-posteriori solution

$$\underset{\boldsymbol{f}_\lambda}{\operatorname{argmax}} \quad \log \pi(\boldsymbol{m}_\lambda | \boldsymbol{f}_\lambda) + \log \pi(\boldsymbol{f}_\lambda) \tag{4}$$

where the conditional probability function $\pi(\boldsymbol{m}_\lambda|\boldsymbol{f}_\lambda)$ is given by the noise model (1) and the probability function $\pi(\boldsymbol{f}_\lambda)$ represents a prior knowledge about the λ-slice $\boldsymbol{f}_\lambda$. To solve (4) for

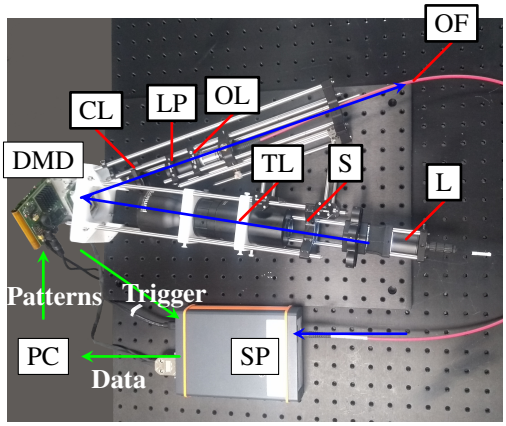

Fig. 2. Acquisition system. Light source (L), sample (S), telecentric lens (TL), digital micromirror device (DMD), bi-convex lens (CL), achromatic lens pair (LP), objective lens (OL), optical fiber (OF), spectrometer (SP), and instrumentation computer (PC). The green arrows indicate the communication workflow between the computer, the DMD, and the spectrometer. The blue arrows indicate the light path.

unknown $\pi(\boldsymbol{f}_\lambda)$, we adopt the deep expectation-maximization network (EM-Net) [21]. Setting the number of iterations to I, the EM-Net $\mathcal{G}_\theta^{(I)}(\boldsymbol{m}_\lambda)$ computes recursively for $0 \leq i \leq I-1$

$$\bar{\boldsymbol{f}}_\lambda^{(i)} = \boldsymbol{f}_\lambda^{(i)} + \mathcal{G}_{\mathrm{dc}}(\boldsymbol{m}_\lambda - \boldsymbol{H}_\downarrow \boldsymbol{f}_\lambda^{(i)}) \tag{5a}$$

$$\boldsymbol{f}_\lambda^{(i+1)} = \mathcal{D}_\theta(\bar{\boldsymbol{f}}_\lambda^{(i)}) \tag{5b}$$

where $\mathcal{G}_{\mathrm{dc}}$ represents Gaussian denoised completion and $\mathcal{D}_\theta$ represents a convolutional neural network with parameters θ. The Gaussian denoised completion corresponds to the solution of (4) under Gaussian assumptions. It can be computed analytically as

$$\mathcal{G}_{\mathrm{dc}}(\boldsymbol{m}) = \frac{1}{N}\boldsymbol{H}_\downarrow^\top \begin{bmatrix} \boldsymbol{I}_M \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1} \end{bmatrix} \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}+\boldsymbol{\Sigma}_1)^{-1}\boldsymbol{m} \tag{6}$$

where $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_{21}$ are blocks of the covariance matrix of $\boldsymbol{Hf}$ and $\boldsymbol{\Sigma}$ is the noise covariance, which can be estimated as detailed in [20]. We optimize the parameters of the convolutional neural network in a supervised manner

$$\operatorname*{argmin}_{\boldsymbol{\theta}} \sum_\ell \|\boldsymbol{f}^{(\ell)} - \mathcal{G}_\theta^{(I)}(\boldsymbol{m}^{(\ell)})\|^2 \tag{7}$$

where $\{\boldsymbol{f}^{(\ell)} \in \mathbb{R}^N\}_{1\leq\ell\leq L}$ is an image database and $\{\boldsymbol{m}^{(\ell)} \in \mathbb{R}^K\}_{1\leq\ell\leq L}$ are the associated measurements computed according to (1) and (2).

2.3. Experimental setup

Our setup, depicted in Fig. 2, is composed of an illumination arm, a DMD, and a light collection arm. The illumination arm is composed of a white LED lamp (Thorlabs LIUCWHA) and a bi-telecentric lens system (Edmund Optics TECHSPEC® Large Format Telecentric 62902, magnification 0.9x) that forms the image of the object in the active plane of a DMD (ViALUX GmbH DLP V-700, 1024 x 768 micromirrors, 13.7 µm pitch). The DMD is made of a matrix

of microscopic mirrors that can be individually tilted to either +24° (ON state) or −24° (OFF state) according to spatial light patterns. The light collection arm, placed at +24° with respect to illumination arm, holds a 35 mm focal length bi-convex lens, a MAP104040-B Matched Achromatic Lens Pair (both focal lengths are 40 mm), and an objective lens (x20, NA=0.35) that focuses light at the entrance of an optical fiber (1500 µm core diameter, NA = 0.39, FT1500 UMT) connected to a compact spectrometer (Avantes AvaSpec-ULS2048CL-EVO, $\Lambda = 2048$ spectral channels, 515–750 nm, entrance slit of 200 µm, 1200 lines/mm grating). The setup, which is supported by a cage system, is lightweight and transportable.

2.4. Acquisition software

After initialization of the DMD and the spectrometer, a sequence of patterns is uploaded into DMD memory, the acquisition is based on the external trigger signal provided by the DMD (see the execution diagrams of Figure 3). Each pattern is displayed on the DMD during a given illumination time and the external trigger is sent to the spectrometer for synchronization. The integration time of the spectrometer is chosen as equal to the illumination time. A dead time of 44 µs, referred to as dark phase, is necessary for the DMD to get the micromirrors tilted according to the next pattern. Another dead time of 356 µs is necessary for the spectrometer to flush its buffer and prepare a new acquisition. This leads to an acquisition time per pattern equal to the spectrometer integration time plus its dead time. The total time for the acquisition of a hypercube is

$$T = 2K(\Delta t + \delta t) \tag{8}$$

where Δt represents the integration time and δt represents the dead time of the spectrometer. While the integration time can be chosen by the user, the dead time is imposed by the spectrometer. Note that the dead time of the spectrometer is much longer that the smallest illumination time allowed by the DMD which cannot be operated at its maximum frequency (22 kHz corresponding to 45 µs).

Our acquisitions are typically made with an integration time of 1 ms. Therefore, the fully sampled acquisition of an image of $N = 64 \times 64$ pixels requires $2K = 2N = 8{,}192$ patterns $\times$ 1.4 ms $\approx$ 11.5 s. During acquisition, the spectra are stored in the spectrometer's internal memory and are transferred to the computer via a callback function, which allows other tasks such as image reconstruction to be run in parallel. The acquisition workflow is implemented as an open source Python package named Single-Pixel Acquisition Software (SPAS) [26], which requires the ALP4lib package [27] for DMD control and MSL-Equipment [28] for spectrometer control.

3. Experiments

3.1. Experimental data

We image four objects illuminated in transmission mode with the LIUCWHA LED lamp: a cat image from the STL-10 test set printed on a plastic sheet, on which we superimpose a linear variable filter (Ocean Optics, LVF-HL, see Fig. 4); the Siemens star resolution target (Thorlabs, R1L1S2P, see Fig. 5); the USAF resolution target (Edmund, USAF 1951 38256, see Fig. 5); and a tomato slice (see Fig. 8). We also image a Mercury-Argon spectral calibration lamp (Ocean Optics HG-1 with characteristic peaks at 546, 577, 579, 697, 707, 727, and 738 nm) directly, i.e., with no object. Whatever the imaging configuration, we acquire all the patterns of a 64×64 Hadamard basis, resulting in a total of $M = N = 4{,}096$ Hadamard patterns split into 8,192 positive and negative patterns. We also consider accelerated acquisition for which only $M <$ 4,096 patterns are taken. The fully sampled datasets can be downsampled *a posteriori* to simulate an accelerated acquisition with different acceleration factors.

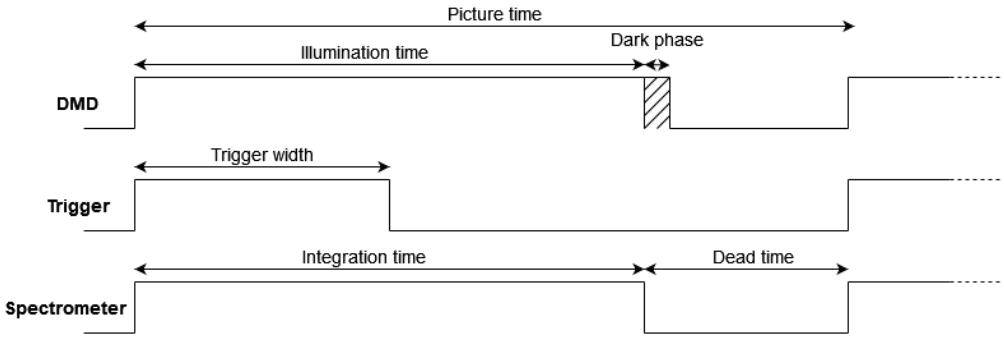

Fig. 3. Synchronization between the digital micromirror device (DMD) and the spectrometer. The external trigger is generated by the DMD (master) and exploited by the spectrometer (slave).

Table 1. Integration times for the different zooms and samples. Times are given in ms/patterns; 'n.a.' indicates that a dataset is not available.

zoom	×1	×2	×3	×4	×6	×12
pixel size (µm)	182.4	91.2	60.8	45.6	30.4	15.2
STL-10 cat	1	4	9	16	36	144
Siemens star	1	4	9	16	36	144
USAF	1	4	9	16	36	144
Hg-Ar lamp	17	n.a.	n.a.	n.a.	n.a.	n.a.
Tomato slice	1	4	9	16	36	144

Typically, the 64×64 Hadamard patterns are resized to fill the largest square region possible on the DMD, which corresponds to 768×768 micromirrors. In this case, each pixel of the Hadamard patterns corresponds to 12×12 micromirrors. However, it is also possible to display the patterns on smaller fields of view, which acts as a hardware zoom that is independent of the optical components of the acquisition setup. We consider six DMD-based hardware zooms: ×1, ×2, ×3, ×4, ×6, and ×12, which correspond to patterns with a pixel size of 12, 6, 4, 3, 2, and 1 micromirrors, respectively. The higher the zoom factor, the lower the photon counts. To get measurements with similar signal-to-noise ratios, we chose the integration time depending on the zoom, as indicated in Table 1. For the ×1 zoom, we set the integration time to 1 ms/pattern for the STL-10 cat and 17 ms/pattern for the spectral calibration source. For the Siemens star and USAF targets that we image at zooms ×1, ×2, ×4, ×6, and ×12, we choose integration times of 1, 4, 16, 36, and 144 ms/pattern, respectively. For the tomato slice, we choose 4 ms/pattern for ×2 zoom and 144 ms/pattern for the ×12 zoom.

Following the procedure described in [24], we estimate the noise parameters as follows: $\mu_{\text{dark}} = 739$ counts, $g = 0.77$ counts/electron, and $\sigma_{\text{dark}} = 17$ counts.

3.2. Training of the EM-Net

We train our network with measurements that we simulate using the STL-10 database [29], with $L = 105,000$ images that correspond to the 'unlabeled' and 'train' subsets. The original 96×96 pixel images are resized to 64×64 pixels using a bicubic transform and are normalized between -1 and 1. As in [20, 21], we choose the light patterns $\boldsymbol{P}$ as 2D Walsh Hadamard functions,

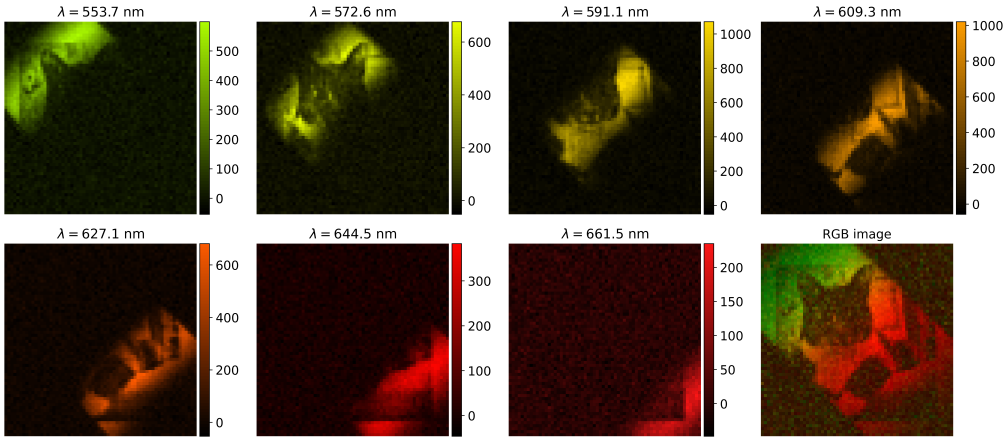

Fig. 4. STL-10 cat hypercube acquisition with a linear variable filter. The full hypercube is binned spectrally for display (7 bins in the range 544–670 nm, bin widths ∼19 nm, central wavelengths: 553.7, 572.6, 591.1, 609.3, 627.1, 644.5, and 661.5 nm). The colorbars show intensities in counts/pixel. The image on the bottom right is an RGB representation of the full hypercube. Acquisition: $K = N = 4{,}906$ patterns, ×1 zoom, integration time of 1 ms/pattern; reconstruction by means of (3).

which we undersample by retaining the patterns that lead to the coefficients with the largest variance [30]. We implement the EM-Net using Pytorch [31] (version 1.10.1; cuda V11.5.50). The image domain denoiser $\mathcal{D}_\theta$ has four convolutional layers, with each layer separated by a ReLU and batch normalization layer. The first has a kernel size of 9 and a depth of 64, the second has a kernel size of 1 and a depth of 64, the third has a kernel size of 3 and a depth of 64, and the final one has a kernel size of 5 and a depth of 1. We initialize our EM-Net using $\boldsymbol{f}_\lambda^{(0)} = \mathbf{0}$ and train it by solving (7) using the ADAM optimizer [32], with an initial learning rate of 10^{-3}, which is halved every 10 epochs, for a maximum of 100 epochs. The training phase took about 320 minutes on a Tesla V100-SXM2 in the case $K = 512$ and was stopped early at 22 iterations after it had reached convergence.

4. Results

4.1. An acquisition example

To show the spatial and spectral capabilities of our hyperspectral camera, we first consider the cat object with a linear variable filter. The integration time was set at 1 ms/pattern leading to a total acquisition time of 11.5 s. Figure 4 shows the full hypercube to which we apply spectral binning to facilitate its display. We compute 7 bins within the 544–670 nm range with a bin width of ∼19 nm (central wavelengths: 553.7, 572.6, 591.1, 609.3, 627.1, 644.5, and 661.5 nm). We also provide an RGB representation of the full hypercube by application of CIE (International Commission on Illumination 1931) color matching functions. These functions model the chromatic response of the three types of cone cells in the human eye [33].

Each bin displays a different band pass window that is selected by the linear filter. As expected, the band pass window translates diagonally within the field of view, from the top left corner to the bottom right corner, as the central wavelength increases. This is also visible on the RGB representation that displays the color palette starting with the green color in the top left corner to the red color in the bottom right corner. The STL-10 cat is visible in the background.

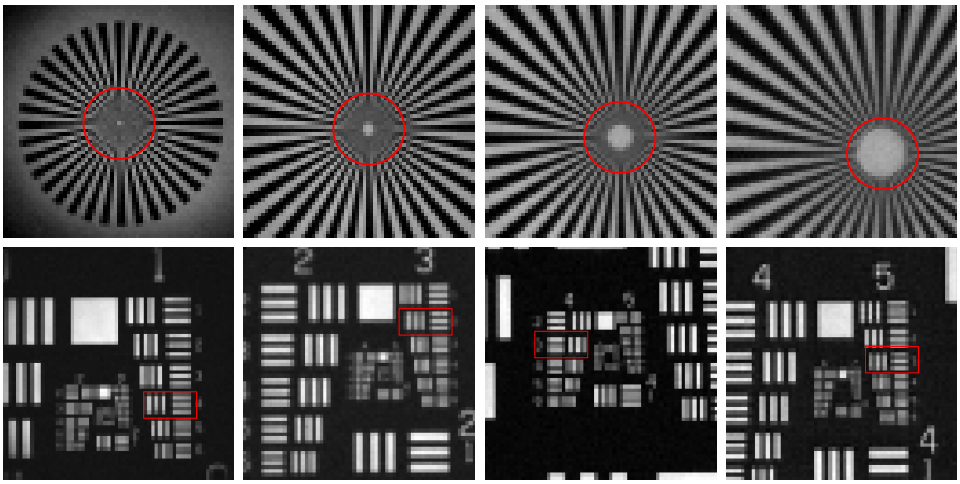

Fig. 5. Resolution targets acquired with different zooms. Top row: Siemens star; bottom row: USAF. Zoom increases from left to right: ×1, ×3, ×6, and ×12; $K = N = 4{,}096$ patterns; the integration time increases with the zoom: 1, 9, 36, and 144 ms/pattern, from left to right. All hypercubes are reconstructed using (3). The displayed images are obtained by summing the hypercubes in the 550-590 nm range.

4.2. Spatial resolution and DMD-based zoom

We evaluate the spatial resolution of our system by imaging two calibrated resolution targets: the Siemens star and the USAF target. The Siemens star is composed of 36 black bars distributed around 360°. The USAF target is composed of bar groups with decreasing bar spacing and length. In Fig. 5, we display the images obtained for both targets at four different zooms (×1, ×3, ×6, and ×12), after summation in the 550–590 nm range. The integration time was set at 1, 9, 36, and 144 ms/pattern respectively leading to a total acquisition time of 11.5, 77, 298.2, and 1,183 s respectively.

For both targets, we first establish the spatial resolution in pixels. Then, we convert it to line pairs per millimeter (lp/mm) to account for the optical magnification. For the Siemens star, we determine the system resolution as the smallest radius of a circular profile for which consecutive black bars appear to touch. For the USAF, we determine the system resolution as the smallest distinguishable bar group vertically and horizontally [34]. We also report the theoretical spatial resolution computed as $1/(2\Delta x)$, where Δx is the image pixel size in millimeters. The image pixel size depends linearly on the zoom, given the DMD pixel size and the telecentric lens magnification. We obtain 182.4, 91.2, 60.8, 45.6, 30.4, and 15.2 µm for ×1, ×2, ×3, ×4, ×6, and ×12 zooms, respectively.

In Fig. 6, we plot the spatial resolution as a function of the zoom, considering six different zooms that correspond to six independent acquisitions. We observe that spatial resolutions obtained from both the USAF and Siemens star targets are in good agreement with theoretical values computed from the pixel size only. This indicates that our system is limited only by the pixel size and that the DMD-based hardware zoom is not associated with undesirable blur.

4.3. Spectral resolution at different locations

We evaluate the spectral resolution by imaging a spectral calibration lamp positioned in the object plane. We place the lamp at three different positions to create light spots at different locations in the field of view and acquire a hypercube for each spot. Figure 7 shows the superposition of the

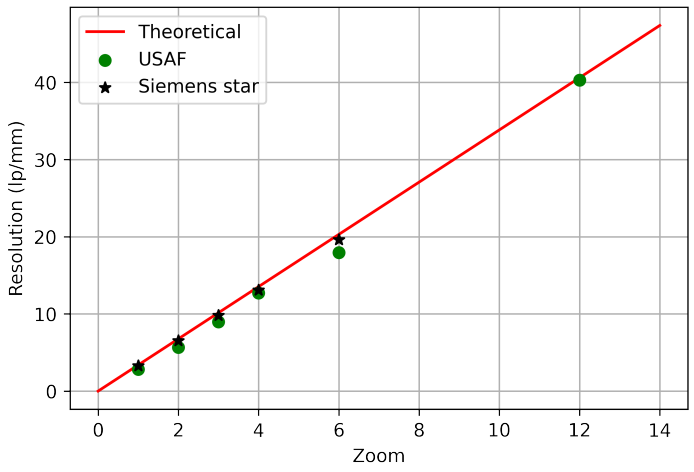

Fig. 6. Spatial resolution as a function of the zoom. The red line is the theoretical resolution calculated from the pixel size; the green dots represent the resolution measured from the USAF target; the black stars represent the resolution measured from the Siemens star target. The spatial resolution is given in line pairs/millimeters (lp/mm) and evaluated from the images displayed in Fig. 5.

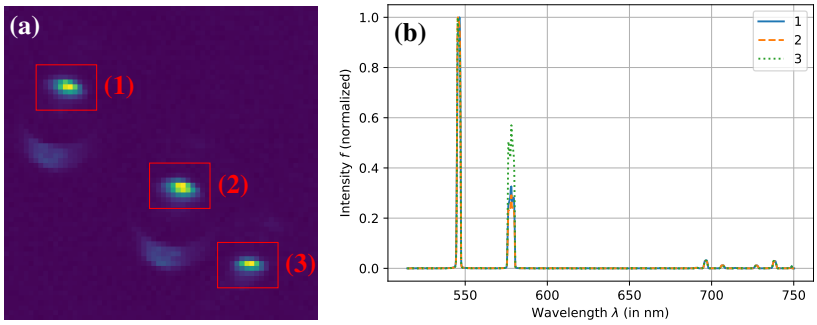

Fig. 7. Spectral resolution at different spatial locations in the field of view. (a) Image of a three-spot calibration source. (b) Spectra of each of the light spots indicated in (a). Acquisition time: 17.4 ms/pattern. The image in (a) is obtained by summing the hypercube along the spectral dimension. The spectra in (b) are obtained by summing all pixels within the red rectangles displayed in (a). These results confirm that the spectral resolution of our device is directly given by the spectral resolution of the spectrometer and that the spectral response of our system is spatially invariant.

three lamp spots ($2K$ = 8,192 patterns per acquisition, integration time Δt = 17.4 ms/pattern). For all three acquisitions we observe a central bright spot corresponding to the position of the light source (Fig. 7(a)). For each spot, we sum the contributions of all pixels in the respective red rectangles, obtaining the spectra indicated in Fig. 7(b).

For the three acquisitions, we recover a spectrum that consists of the emission lines of mercury ($\lambda < 650$ nm) and argon ($\lambda > 650$ nm). In the following, we consider the peaks at 546, 697, 707, 727, and 738 nm. We measure the full width at half maximum of all peaks, for all spot locations positions, and obtain spectral resolutions between 2.15 nm and 2.30 nm. These spectral resolutions are in excellent agreement with the theoretical spectral resolution of the spectrometer that is 2.3 nm, confirming that the spectral resolution of our device is directly given by the spectral resolution of the spectrometer. We observe no spectral degradation that originates from components before the spectrometer (e.g., DMD or focusing optics). Note that the peak at 578 nm results from the observation of the mercury emission doublet at 577 and 579 nm, which cannot be resolved. We also find that the amplitude of the different peaks, except the doublet, are the same for all spot locations, which indicates that the spectral response of our system is spatially invariant.

4.4. Increase of imaging speed via subsampling

We evaluate our ability to reconstruct images from accelerated acquisition considering three samples: the Siemens star target, the USAF target, and the tomato slice. The Siemens star and USAF targets are imaged using the ×12 zoom, while the tomato slice is imaged using both the ×2 and ×12 zooms. For each case, we consider three acceleration factors 1:2 (K = 2,048 measurements), 1:4 (K = 1,024 measurements), and 1:8 (K = 512 measurements). We also reconstruct the hypercubes with no acceleration (K = 4,096 measurements). The images obtained at λ = 579 nm are displayed in Fig. 8. As expected, accelerated acquisitions lead to a loss of spatial resolution, which can be evaluated from the reconstructions of the resolution targets (see first and second row of Fig. 8). The higher the acceleration factor, the higher the loss. For the Siemens star, the degradation of the spatial resolution appears as a blurred region in the center of the target, where high spatial frequency structures are present. We also observe this effect in the tomato slice images (see third and fourth row of Fig. 8). However, as fewer high frequencies are present, the degradation appears relatively limited, even for acceleration factors as high as 1:4 or 1:8.

5. Discussion

A key advantage of our computational design over previous work is to maintain a high spectral resolution. Moreover, its price is significantly lower than currently available hyperspectral cameras with the same spectral resolution. Our system acquires a $64 \times 64 \times 2048$ hypercube with a spectral resolution of 2.3 nm, while the spatial resolution can be adjusted between 182 µm and 15 µm using a DMD-based hardware zoom that can achieve a ×12 magnification with no modification of the optical components. As for optical zoom, the higher the magnification, the lower the photon flux. To account for this effect, the images at higher zooms have been acquired for longer duration (see Fig. 5), with a scaling factor equivalent to the zoom squared.

There are different strategies to limit the total time T given by (8). The first strategy consists in reducing the acquisition time Δt. Setting Δt to 9 µs/pattern, which corresponds to the shortest acquisition time allowed by the the spectrometer, we obtain a total acquisition time of 3 s. However, the spectrometer imposes a dead time δt of 356 µs during which no signal is acquired. For an integration time of 1 ms/pattern, this represents a waste of $356/(356 + 1000) \approx 26\%$ of the total acquisition time. For an integration time of 9 µs/pattern, the waste increases to $356/(356 + 9) \approx 97$ %, i.e., most of the total acquisition time is lost. As a compromise, the shortest integration time that we consider is 1 ms/pattern, leading to a total time of 11.5 s. In the

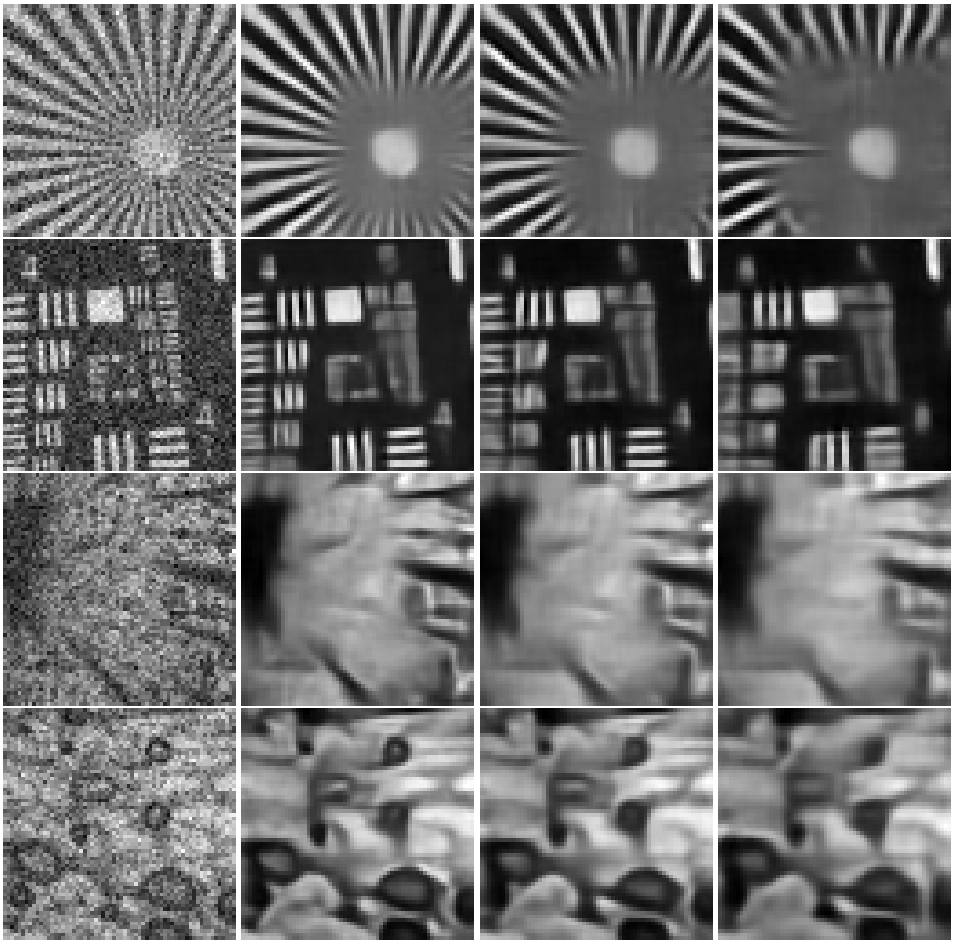

Fig. 8. Accelerated acquisitions. First row: Siemens star ×12 zoom; second row: USAF ×12 zoom; third row: tomato slice ×2 zoom; fourth row: tomato slice ×12 zoom. First column: no acceleration factor, $K = 4{,}096$ measurements; second column: acceleration factor 1:2, $K = 2{,}048$ measurements; third column: acceleration factor 1:4, $K = 1{,}024$ measurements; fourth column: acceleration factor 1:8, $K = 512$ measurements. The fully sampled hypercubes $K = N = 4{,}096$ are reconstructed using (3), while the accelerated acquisitions $K < N$ are reconstructed using the deep expectation-maximization network (EM-Net) defined by (5). All images correspond to the spectral channel $\lambda = 579$ nm.

future, spectrometers with shorter dead times could mitigate this issue. Assuming negligible dead times, i.e., $\delta t \ll \Delta t$, an acquisition with $\Delta t = 1$ ms would drop from 11.5 s to 8.6 s, while the fastest acquisition with $\Delta t = 45$ µs would drop from 3.3 s to 0.4 s. Another strategy to reduce the total acquisition time consists in limiting the number of patterns $2K$ uploaded onto the DMD. The total acquisition time of so-called accelerated acquisition depends directly on the acceleration factor (e.g., $11.5/2 \approx 5.75$ s considering only half of the patterns). This acceleration comes at the cost of spatial resolution reduction, as illustrated in Fig. 8. The acceptable acceleration factor depends on the frequency content of the scene. While a 2-fold acceleration may be already excessive for sharp or highly structured objects, an acceleration up to 8-fold may be acceptable for smoother objects. It is important to note that our algorithm reconstructs each λ-slice of the hypercube independently; however, it could be beneficial to exploit the spatio-spectral redundancy in order to jointly reconstruct several λ-slices (see for instance [35]). This could enable the same spatial resolution to be achieved with higher acceleration factors. Moreover, the determination of the best subset of patterns remains an open problem and the subject of an active research, including different fields such as magnetic resonance imaging.

Another limitation of our imaging system is its spatial resolution. For applications where the imaging speed is not the limiting factor, this can be alleviated by increasing the number of patterns. However, the amount of DMD memory currently available does not allow more that 43,690 binary patterns to be stored. Therefore, our setup can acquire hypercubes with 128×128 pixels (32,768 patterns required) but not 256×256 pixels (131,072 patterns required). This issue can be mitigated by considering accelerated acquisitions (e.g., 3-fold acceleration for 256×256 pixels). In the future, DMDs with more memory could remove this barrier.

One challenge is to maximize the light throughput. There is an inherent trade-off between light collection by the optical fiber and by the spectrometer. Due to etendue conservation, more light can be focused at the entrance of large core diameter (e.g., 1500 µm) fibers. However, due to the finite size of the entrance slit of the spectrometer (e.g., 200 µm), increasing the size of the optical fiber, increases proportionally the number of rejected photons. We have chosen the optical components, distances, magnifications, and optical fiber in order to maximize the signal intensity measured by the spectrometer at $\times 1$ zoom.

6. Conclusion

We propose a hyperspectral imager capable of acquiring a $64 \times 64 \times 2048$ hypercube with a spectral resolution of 2.3 nm. The spatial resolution can be adjusted between 182.4 µm and 15.2 µm using a digital zoom. Setting the integration time to 1 ms per pattern, the total acquisition time for a single hypercube is less that 12 s. The total acquisition time can be reduced to 2.6 s by reducing the integration time to 9 µs. Such small integration times should be reserved for high intensity signals as they lead to a 97 % loss due to dead times. While maintaining the integration time of 1 ms per pattern, the acquisition time can be accelerated by uploading fewer patterns onto the DMD. For several acceleration factors, we demonstrate that the slices of the hypercube can be reconstructed independently using a deep EM-Net. 8-fold acceleration leads to a 1.4 s acquisition and reconstructions with moderate degradation for low frequency images. In future work, we will consider coupling our hyperspectral imager with a standard camera to improve both the spatial resolution and imaging speed.

Funding. This study was supported by the French National Research Agency (ANR), under Grant ANR-17-CE19-0003 (ARMONI Project), and performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon.

Acknowledgments. This material is based on work carried out at the PILoT facility (PILoT, INSA LYON, Bât. Blaise Pascal, 7 Av. Jean Capelle 69621 Villeurbanne).

Disclosures. The authors declare that they have no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Ref. [36].

References

1. B. Boldrini, W. Kessler, K. Rebner, and R. W. Kessler, "Hyperspectral Imaging: A Review of Best Practice, Performance and Pitfalls for in-line and on-line Applications," J. Near Infrared Spectrosc. **20**, 483–508 (2012).
2. Q. Li, X. He, Y. Wang, H. Liu, D. Xu, and F. Guo, "Review of spectral imaging technology in biomedical engineering: Achievements and challenges," J. Biomed. Opt. **18**, 100901 (2013).
3. A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the Earth: A personal view," Remote. Sens. Environ. **113**, S5–S16 (2009).
4. J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy," Nat. Methods **2**, 910–919 (2005).
5. N. A. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," Opt. Eng. **52**, 090901 (2013).
6. J. P. Maillard, L. Drissen, F. Grandmont, and S. Thibault, "Integral wide-field spectroscopy in astronomy: The Imaging FTS solution," Exp. Astron. **35**, 527–559 (2013).
7. J. G. Dwight and T. S. Tkaczyk, "Lenslet array tunable snapshot imaging spectrometer (LATIS) for hyperspectral fluorescence microscopy," Biomed. Opt. Express **8**, 1950–1964 (2017).
8. X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational Snapshot Multispectral Cameras: Toward dynamic capture of the spectral world," IEEE Signal Process. Mag. **33**, 95–108 (2016).
9. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive Coded Aperture Spectral Imaging: An Introduction," IEEE Signal Process. Mag. **31**, 105–115 (2014).
10. Y. Oiknine, I. August, V. Farber, D. Gedalin, and A. Stern, "Compressive Sensing Hyperspectral Imaging by Spectral Multiplexing with Liquid Crystal," J. Imaging **5**, 3 (2019).
11. K. Monakhova, K. Yanny, N. Aggarwal, and L. Waller, "Spectral DiffuserCam: Lensless snapshot hyperspectral imaging with a spectral filter array," Optica **7**, 1298 (2020).
12. F. Soldevila, J. Dong, E. Tajahuerce, S. Gigan, and H. B. de Aguiar, "Fast compressive Raman bio-imaging via matrix completion," Optica **6**, 341 (2019).
13. M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," Nat. Photonics **13**, 13–20 (2019).
14. G. M. Gibson, G. M. Gibson, S. D. Johnson, S. D. Johnson, M. J. Padgett, and M. J. Padgett, "Single-pixel imaging 12 years on: A review," Opt. Express **28**, 28190–28208 (2020).
15. F. Rousset, N. Ducros, A. Farina, G. Valentini, C. D'Andrea, and F. Peyrin, "Adaptive basis scan by wavelet prediction for single-pixel imaging," IEEE Transactions on Comput. Imaging **3**, 36–46 (2017).
16. Q. Pian, R. Yao, N. Sinsuebphon, and X. Intes, "Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging," Nat. Photonics **11**, 411–414 (2017).
17. E. Hahamovich, S. Monin, Y. Hazan, and A. Rosenthal, "Single pixel imaging at megahertz switching rates via cyclic Hadamard masks," Nat. Commun. **12**, 4516 (2021).
18. R. I. Stantchev, X. Yu, T. Blu, and E. Pickwell-MacPherson, "Real-time terahertz imaging with a single-pixel detector," Nat. Commun. **11**, 2535 (2020).
19. C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," Sci. Reports **8**, 2369 (2018).
20. A. Lorente Mur, P. Leclerc, F. Peyrin, N. Ducros, and N. Ducros, "Single-pixel image reconstruction from experimental data using neural networks," Opt. Express **29**, 17097–17110 (2021).
21. A. Lorente Mur, P. Bataille, F. Peyrin, and N. Ducros, "Deep Expectation-Maximization For Image Reconstruction From Under-Sampled Poisson Data," in 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), (IEEE, Nice, France, 2021), pp. 1535–1539.
22. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nat. Rev. Phys. **3**, 422–440 (2021).
23. "Openspyrit/spyrit," openspyrit (2021).
24. E. M. V. Association, "EMVA standard 1288, standard for characterization of image sensors and cameras," Release **3**, 29 (2016).
25. A. Lorente Mur, M. Ochoa, J. Cohen, X. Intes, and N. Ducros, "Handling negative patterns for fast single-pixel lifetime imaging," in Molecular-Guided Surgery: Molecules, Devices, and Applications V, B. W. Pogue and S. Gioux, eds. (SPIE, San Francisco, United States, 2019), p. 9.
26. G. Beneti Martins, L. Mahieu-Williame, and N. Ducros, "Single-pixel acquisition software version 1.0," (2021).
27. S. M. Popoff and M. W. Matthès, "ALP4lib: A Python wrapper for the Vialux ALP-4 controller suite to control DMDs," Zenodo (2020).
28. Measurement Standards Laboratory of New Zealand, "MSL-Equipment," .
29. A. Coates, A. Ng, and H. Lee, "An Analysis of Single-Layer Networks in Unsupervised Feature Learning," in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, (JMLR Workshop and Conference Proceedings, 2011), pp. 215–223.
30. L. Baldassarre, Y.-H. Li, J. Scarlett, B. Gözcü, I. Bogunovic, and V. Cevher, "Learning-Based Compressive Subsampling," IEEE J. Sel. Top. Signal Process. **10**, 809–822 (2016).
31. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai,

and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems 32, H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, eds. (Curran Associates, Inc., 2019), pp. 8024–8035.
32. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, Y. Bengio and Y. LeCun, eds. (2015).
33. H. S. Fairman, M. H. Brill, and H. Hemmendinger, "How the CIE 1931 color-matching functions were derived from Wright-Guild data," Color. Res. & Appl. **22**, 11–23 (1997).
34. A. Orych, "REVIEW OF METHODS FOR DETERMINING THE SPATIAL RESOLUTION OF UAV SENSORS," The Int. Arch. Photogramm. Remote. Sens. Spatial Inf. Sci. **XL-1/W4**, 391–395 (2015).
35. V. Pronina, A. Lorente Mur, J. F. P. J. Abascal, F. Peyrin, D. V. Dylov, and N. Ducros, "3D denoised completion network for deep single-pixel reconstruction of hyperspectral images," Opt. Express **29**, 39559 (2021).
36. L. Mahieu-Williame and N. Ducros, "Single-Pixel Camera Datasets (version 2.0)," (2022).