Park and Gao *PhotoniX (2024)5:37*
https://doi.org/10.1186/s43074-024-00152-5

PhotoniX

RESEARCH Open Access

Cascaded compressed-sensing single-pixel camera for high-dimensional optical imaging

Jongchan Park[1] and Liang Gao[1*]

*Correspondence:
gaol@ucla.edu

[1] Department of Bioengineering, University of California, Los Angeles, CA 90095, USA

Abstract

Single-pixel detectors are popular devices in optical sciences because of their fast temporal response, high sensitivity, and low cost. However, when being used for imaging, they face a fundamental challenge in acquiring high-dimensional information of an optical field because they are essentially zero-dimensional sensors and measure only the light intensity. To address this problem, we developed a cascaded compressed-sensing single-pixel camera, which decomposes the measurement into multiple stages, sequentially reducing the dimensionality of the data from a high-dimensional space to zero dimension. This measurement scheme allows us to exploit the compressibility of a natural scene in multiple domains, leading to highly efficient data acquisition. We demonstrated our method in several demanding applications, including enabling tunable single-pixel full-waveform hyperspectral light detection and ranging (LIDAR) for the first time.

Keywords: Single-pixel camera, Compressed sensing, LiDAR, Hyperspectral imaging

Introduction

Although century-old, single-pixel detectors are still prevalent in optical sciences because of their fast temporal response, high sensitivity over a broad spectrum, low dark noise, and low cost [1, 2]. However, when being used for imaging, they are fundamentally limited by a "*dimensionality gap*"—while an optical field is characterized by a multidimensional plenoptic function [3] with rich information, the single-pixel detector measures only the light intensity, leading to a zero-dimensional (0D) image.

The convergence of recent advances in spatial light modulating technologies [4] and computational imaging shows great promise in bridging this gap. The resultant "single-pixel camera" has been demonstrated in myriad imaging applications [1, 2, 5–13]. Moreover, using compressed sensing in single-pixel cameras [6, 14, 15] further allows capturing light information with a sampling rate much lower than the Nyquist condition, leading to a more efficient data acquisition than conventional pixel-array cameras. The reduced sampling rate is particularly crucial for high-dimensional (high-D) imaging systems that inherently carry a heavy dataload. Nonetheless, the adoption of compressed single-pixel cameras in multidimensional optical imaging is hindered by two significant technical challenges: the choice of an appropriate compressed

sampling scheme and the optical transformation of the data from a high-D space to a low-dimensional (low-D) space.

On the one hand, when choosing an appropriate sampling scheme for compressed imaging, the most common choice is to use a Hadamard basis. The Hadamard basis measurement, derived from a Hadamard matrix [16, 17] consisting of elements of $+1$ and -1, can be readily implemented in an optical system with a digital micromirror device (DMD) [17, 18], which modulates the light in a binary manner. A typical DMD-based compressed imaging system encodes the scene with a sequence of Hadamard patterns and measures the integrated signals using a single-pixel detector at each pattern. The image can then be reconstructed by correlating the encoding patterns with the signals obtained. However, when capturing an image at a sub-Nyquist sampling rate, that is, the image's pixel count is much greater than the number of acquired samples, choosing an appropriate subset of the Hadamard basis becomes critical as it is often object-dependent [19, 20]. Although regularization algorithms [21, 22] may remedy the result, they are ineffective in processing highly compressed measurements [23].

On the other hand, capturing optical information beyond a two-dimensional (2D) intensity map—such as in hyperspectral [24] or volumetric imaging [3]—with a single-pixel detector is nontrivial because most current spatial light modulators are in a 2D format. For multidimensional optical imaging, a high-D optical field must be first transformed to a low-D space so it can be modulated by a spatial light modulator [25–27]. However, the number of voxels that can be captured from the high-D scene is limited by the pixel count of the 2D spatial light modulator, leading to a trade-off between resolutions along the spatial and multidimensional axes. An alternative way is to perform imaging through temporal multiplexing. For example, by coupling a Fourier transform imaging spectrometer to a single-pixel camera, Jin et al. demonstrated compressed single-pixel hyperspectral imaging with a broad spectral range. The corresponding voxel size of the hyperspectral datacube is $(256 \times 256 \times 701)$ [28]. However, the trade-in of a single-pixel detector's temporal bandwidth for spectral information makes the device incapable of measuring dynamic light [29], such as time-of-flight signals. Additionally, in volumetric imaging, most systems rely on structured light illumination [10, 30], which limits the ability to image scenes under natural light.

To address these challenges, we develop a new category of cascaded compressed-sensing single-pixel cameras. Rather than relying on a single process to modulate a multidimensional light field, we decompose the compressed measurement into cascaded stages, sequentially reducing the dimensionality of the data from a high-D space to 0D. Such a measurement scheme allows us to exploit the compressibility of an object in multiple domains, which can lead to higher reconstructed image quality. In addition, the voxel size of the high-D scene is not constrained by the pixel count of the spatial light modulator, unlike in the single-step compression method. Moreover, the accessibility to the intermediate image between adjacent measurement stages allows the mapping of high-D optical information to a low-D space through pure optical operations without sacrificing the temporal bandwidth of single-pixel detectors. We demonstrated our method in several formerly intractable applications, including enabling single-pixel full-waveform hyperspectral light detection and ranging (LIDAR) for the first time.

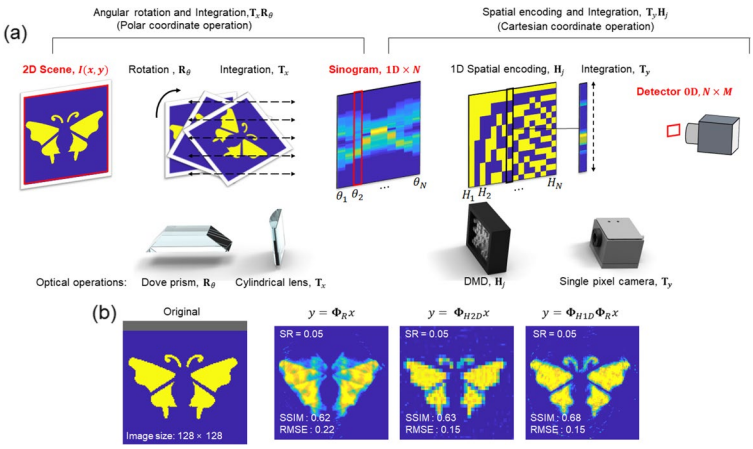

Fig. 1 Principle of cascaded compressed sensing. **a** Illustration of a single-pixel camera with cascaded compression. The 2D scene is rotated axially and integrated along the horizontal direction by a Dove prism and the cylindrical lens. The resultant multiple 1D images at different rotation angles are composed into a sinogram of *en-face* parallel projections, similar to computed tomography. A DMD encodes a binary pattern onto the sinogram and integrates the signals along the horizontal direction. The single-pixel detector captures the resultant zero-dimensional data. **b** Compressed imaging by using the *en-face* parallel projection method ($y = \mathbf{\Phi}_R x$), 2D Hadamard-basis compression method ($y = \mathbf{\Phi}_{H2D} x$), and cascaded compression method ($y = \mathbf{\Phi}_{H1D}\mathbf{\Phi}_R x$). x and y are the incident image and the measurement. $\mathbf{\Phi}_R$, $\mathbf{\Phi}_{H2D}$, and $\mathbf{\Phi}_{H1D}$ are linear sensing matrices of *en-face* projection, 2D Hadamard encoding, and 1D Hadamard encoding, respectively. The sampling ratio is 0.05. SSIM structural similarity; RMSE, root mean squared error

Results

Cascaded compressed measurement

The core idea of our method is to transform light information in a high-D space to 0D through a cascaded process. For example, we show the optical realization of a two-stage system in Fig. 1, where we use two coupled subsystems to sequentially perform compressed measurements in the angular and spatial domains. In Stage I, we rotate the incident image using a Dove prism, followed by optically integrating the signals along the horizontal axis using a cylindrical lens. This transforms the input 2D image into a 1D line, essentially an *en-face* projection (or the radon transformation [31]) of the object [32, 33]. Repeating this operation at different image rotation angles yields a sinogram like in computed tomography [34, 35]. In Stage II, we take the 1D projection image from Stage I as the input and encode it with a Hadamard binary pattern using a spatial light modulator. The resultant light signals are then focused by a lens and spatially integrated at a single-pixel detector. The image formation model of the entire system can be formulated as:

$$E_{j,\theta} = \mathbf{V}_j\left([1 \quad 1 \quad 1 \quad \cdots \quad 1]\mathbf{R}_\theta I(x,y)\right)^{\mathrm{T}} = \mathbf{T}_y\mathbf{V}_j\mathbf{T}_x\mathbf{R}_\theta I(x,y). \tag{1}$$

Here $\mathbf{T}_x$ and $\mathbf{T}_y$ are integration operators along the horizontal and vertical axis, respectively. $\mathbf{V}_j$ is a 1D binary spatial encoding operator, and $\mathbf{V}_j = 1/2(\mathbf{H}_1 + \mathbf{H}_j)$, where $\mathbf{H}_j$ is the j^{th} Hadamard row vector. $I(x,y)$ denotes the incident 2D intensity image of $N \times N$ pixels. $\mathbf{R}_\theta$ is a truncated rotating operator, maintaining the image size. $E_{j,\theta}$ is the light intensity measured by the single-pixel detector. By choosing a subset of angular

projections $\theta = \theta_1, \theta_2, \ldots, \theta_r$ and spatial samplings $j = 1,2,\ldots,k$, we can capture the original 2D image compressively with a sampling ratio (SR) of $(r \times k)/N^2$.

Because optical operations are linear, we reconstruct the scene $I(x, y)$ from the sparsely sampled measurement, $\mathbf{E}$, by solving a linear inverse problem:

$$\text{argmin}_I \|\mathbf{E} - \mathbf{\Phi} I\|_2^2 + \lambda\varphi(I), \tag{2}$$

where $\mathbf{E}$ is the vector representation of $E_{j,\theta}$ of size $1 \times (rk)$, and $\mathbf{\Phi} = \mathbf{T}_y\mathbf{V}_j\mathbf{T}_x\mathbf{R}_\theta$ is the sensing matrix that describes the linear relationship between the signal and the measurement. λ is the regularization parameter, and $\phi(I)$ is the regularization function, where we use l_1-norm to force sparsity.

For high-quality image reconstruction, the system must satisfy two conditions: signal sparsity and incoherence of the measurement [36]. The signal is considered sparse if it can be represented in a few high-valued coefficients in an orthonormal basis as $I = \mathbf{\Psi}$ s, where $\mathbf{\Psi}$ is the sparse basis matrix, and s is a sparse representation of the signal (i.e., $k \ll N^2$, where k is the number of high-valued coefficients of s). To exploit the sparsity of the signal and recover the signal with high fidelity, the measurement matrix $\mathbf{\Phi}$ must be incoherent to the sparse basis matrix $\mathbf{\Psi}$. In other words, the correlation between any two columns in $\mathbf{\Phi}$ and $\mathbf{\Psi}$ must be minimized [37]. In fact, the convex optimization algorithm of l_1-norm minimization is also referred to as a basis-pursuit method [38] that recovers the signal by finding its sparse representation.

However, without a priori knowledge of the signal, finding an optimal sensing matrix (or optimal measurement scheme) mutually incoherent to the sparse basis matrix is a demanding task [39]. For example, in Hadamard-basis compressed imaging, the image can be well reconstructed only when the signal is sparse in its original domain (space) but not in a transformed domain [40]. Also, the image reconstruction quality is susceptible to the choice of the subset of the Hadamard basis at a given sampling rate, especially under a highly under-sampled condition with noises [23]. In contrast, our method completely reshuffles the signals by a series of optical operations, including radon transformation (*en-face* parallel projection) and spatial projection, thereby facilitating the image reconstruction through the basis-pursuit l_1-norm minimization (Fig. 1b).

Numerical simulations and experimental demonstrations

We performed numerical simulations to evaluate our method at a highly sub-Nyquist sampling rate. We chose a standard Shepp-Logan phantom image with 128×128 pixels as the test image (Fig. 2a) and simulated three compressed measurements: *en-face* parallel projection, 2D Hadamard basis compressed imaging, and cascaded compression. The target image was processed using the image formation model of each method, and white random gaussian noises were added to the image with a signal-to-noise ratio of 20 dB. We reconstructed the image using an iterative optimization algorithm (FISTA [41] with a proximal operator of the l_1-norm, see MATLAB code in Supplementary Information). To choose the subset of the Hadamard basis, we used total variation ordering [23] because it typically yields a high-quality image given a low sampling rate, whereas no prior scene information is needed. For the *en-face* projection measurement, the

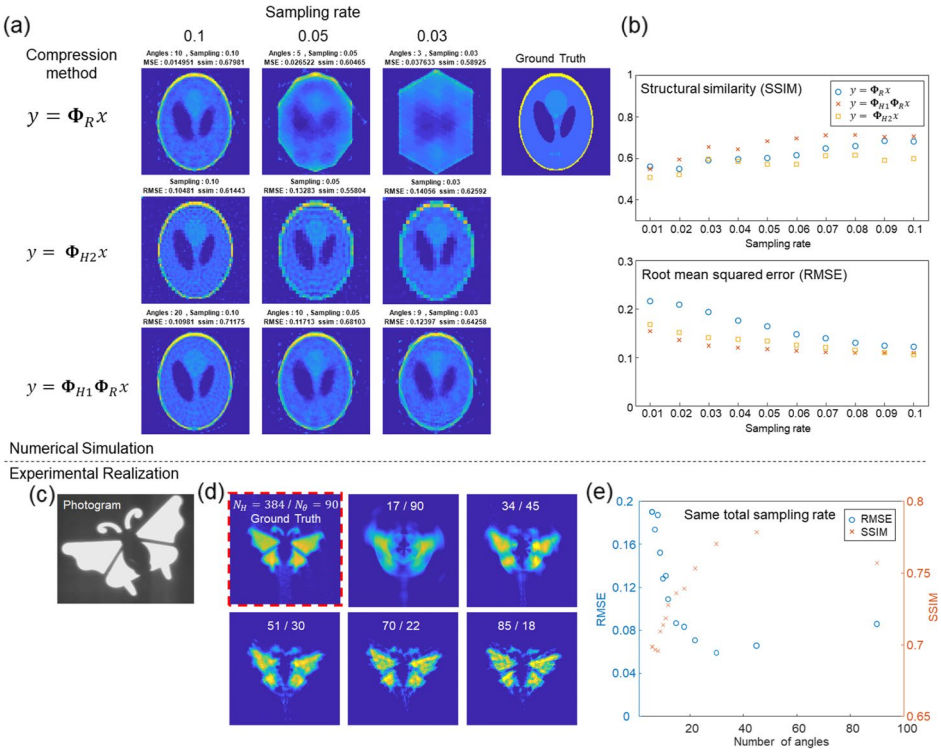

Fig. 2 Image compression with three approaches: 2D Hadamard-basis compression method, ***en-face***parallel projection method (sparse angle radon transformation), and cascaded compression method. **a** Reconstructed images with three different compression methods at a low sampling ratio. **b** RMSE and SSIM of the reconstructed images. **c** Photograph of a target scene. **d** Experimental realization of the cascaded compressed measurement. **e** RMSE and SSIM of the reconstructed scene. N_H and N_θ are the number of Hadmadard patterns and angles, respectively

projection angles are uniformly distributed in an angular range of [0, π] to maximize the information content. For a fair comparison, the sampling ratio is calculated as the sampling rate divided by the pixel count of the imaging field of view. Specifically, the field of view is $128 \times 128 = 16{,}384$ pixels in the Hadamard basis compression method. In the other two methods, the field of view is computed as $\pi \times 64^2 = 12{,}867$, whereas it captures information within only the inscribed circular area of a squared scene.

The reconstructed images were evaluated based on two metrics: structural similarity (SSIM) and root mean squared error (RMSE). Figure 2a shows the reconstructed image from each method. The results indicate that the cascaded compression method provides a lower RMSE and a higher SSIM than the other two methods. In a highly undersampled case (SR ≤ 0.05), the limited number of angles in the *en-face* projection method blurs the image along the angular direction. And the 2D Hadamard basis imaging method is plagued by pixelated artifacts because high-spatial frequency information is not addressable. Although a highly compressed inverse problem may be solved by using state-of-the-art optimization algorithms or deep-learning-based approaches [42], it generally requires extensive computation or fine-tuning of the network with pre-acquired training datasets.

Next, we experimentally demonstrated single-pixel imaging using cascaded compression. A 2D image (Fig. 2c) was sequentially compressed by a two-stage system,

and the light intensity was recorded at each Dove prism rotation angle and Hadamard pattern. The reconstructed image is shown in Fig. 2d, where the image size is approximately $\pi \times r^2 = \pi \times 192^2 \sim 116{,}000$ pixels, and the samples acquired are 1,530. The corresponding sampling rate is SR = 0.013. Despite degraded image quality, the cascaded compression method provides a visually identifiable image with only 1.3% of the sampling rate. While the image quality improves as the overall sampling rate increases (Supplementary Fig. S1), the balance between the number of projection angles and the number of Hadamard patterns at a given sampling rate is critical to the image quality. Figure 2e illustrates the RMSE and SSIM of the reconstructed image compared to the densely sampled image. Noteworthily, there is no such an universal optimal measurement scheme [39] because different images possess varied levels of sparsity. Our cascaded compression method provides an efficient sampling by enabling tunability on the sampling rate at each compression step.

Single-pixel spectral imaging (*x*, *y*, *λ*)

Compared with conventional single-stage systems, cascaded compressed sensing provides unique access to the intermediate image between adjacent measurement stages. Optical de-multiplexing of these images allows high-D light field information mapping to a low-D space, which can be further compressed in the subsequent stages. For example, in the two-stage system presented in Fig. 1a, the input image is first transformed into the 1D line through *en-face* projection. The reduced spatial dimension can be refilled with light information of higher dimensions, such as spectrum or polarization. We demonstrated this framework in hyperspectral imaging using a single-pixel camera (Fig. 3a). After Stage-I transformation, we spectrally dispersed the 1D line image using a diffraction grating, mapping the spectral information (500–650 nm) to the orthogonal spatial axis and restoring the spatial dimensions to 2D. The resultant image is then passed to the Stage II subsystem, where we use DMD for spatial encoding. We binned the DMD pixels along the spectral dispersion direction into ten super-pixels, each modulating a different light wavelength.

To demonstrate hyperspectral imaging, we first calibrated the relation between the spectrum of the incident light and the spatial location of the light mapped onto the DMD. We positioned a pinhole at the object plane of the system and illuminated it with 510 nm, 532 nm, 550 nm, 590 nm, and 630 nm wavelength light (Fig. 3b). On the DMD, the light signals are spectrally dispersed along the horizontal axis and sampled by 360 horizontal pixels. Along the vertical direction, the DMD pixels were grouped into 90 stripes, each having a width of four pixels. We sequentially turned on each stripe on the DMD and recorded the corresponding intensities on the single-pixel detector, similar to a push-broom scanner. Horizontal locations of the spectrally filtered lights on the DMD were founded by interpolations between the horizontal pixel index of the stripes and the measured intensity. The linear relations between the spectrum and the spatial locations of the light are illustrated in Fig. 3c. We calibrated the spectral resolution of the system using a beam with a narrow spectrum (532 nm, FWHM 1 nm). The FWHM of the measured spectrum is approximately 15 DMD pixels that is corresponding to 6.2 nm (Fig. 3d). A total of 360 horizontal pixels of the DMD was used to map the spectral information of

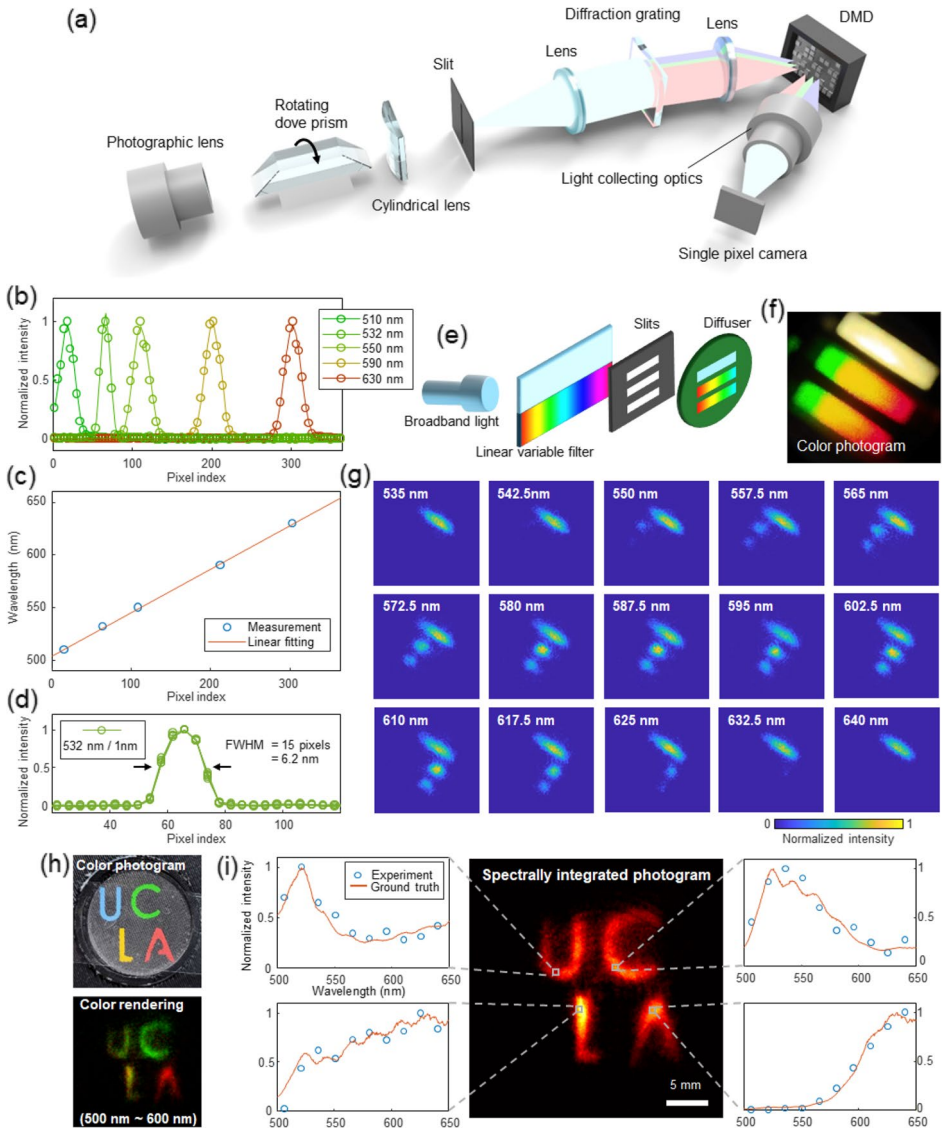

Fig. 3 Single-pixel spectral imaging. **a** System setup. After the *en-face* parallel projection, a 1D vertical line image was spectrally dispersed along the horizontal axis by a diffraction grating and mapped onto the DMD. **b** Spectral calibration of the system. **c** Spectral dispersion curve. **d** Spectral resolution. **e** Imaging target object made with a linear variable filter. **f** Photogram of the target object. **g** Visualization of the image at different wavelengths. **h** Photogram of the target object (top) and color-rendered image based on the spectral irradiance of the scene measured by the single-pixel camera (bottom). Spectral irradiance within the range of 500 nm to 650 nm was used to render. **i** Comparison of spectral irradiance measured by the single-pixel camera and a benchmark fiber spectrometer. The spectrum was spatially integrated over 5 × 5 pixels

the scene within the range of approximately 500 nm to 650 nm. Therefore, the maximum number of resolvable spectral channel in our system is approximately 24.

We tested our method on a colorful object made by using linear variable filter, chrome mask, and diffuser (Fig. 3e). The linear variable filter is a narrow bandpass filter that the spectral transmittance varies along the spatial axis. We back-illuminated the object with a broadband laser and imaged the transmitted light. The chrome mask is consisting of three slits where the the broadband light is directly pass through the top slit and

spectrally filtered light pass through the two bottom slits (Fig. 3f). To capture a hyperspectral datacube of dimensions ($384 \times 384 \times 20$), we used a total of 45 image rotation angles and 192 orthonormal Hadamard patterns in Stage I and II compressed measurements, respectively, leading to an overall SR of 0.075 (the information content is within the inscribed circular area of the 384×384 pixels squared image). The DMD displayed the patterns at 8 kHz, and the total acquisition time was approximately 22 s (Supplementary Table S1 and Supplementary Fig. S2), the corresponding image acquisition speed is approximately at 130,000 voxels/s by utilizing the compressed sensing principle. The voxel rates can be further increased by using spatial and spectral multiplexed measurements [43](Supplementary Fig. S3). The spatial and spectral profile of the object is shown in Fig. 3g.

We quantitatively analysed the accuracy of the spectral measurement by using an object consisting of cellophane tapes and a mask object (Fig. 3h). We back-illuminated the object with a halogen lamp. The spectra of the beam pass through each cellophane tape were obtained with a benchmark fiber spectrometer (STS-VIS-L-25-400-SMA, Ocean Optics) and served as ground truth values. In this measurement, the voxel size of the hyperspectral datacube is ($384 \times 384 \times 10$). The measured spectra are illustrated in Fig. 3i. The RMSE of the normalized spectral difference is 0.092.

It is noteworthy that the sampling rates of our cascaded compressed-sensing camera can be tuned in both spatial and spectral domains, allowing a tailored compressed measurement for a given scene. In the spatial domain, the sampling rate can be tuned by choosing the number of image rotation angles and encoding patterns. While in the spectral domain, because the spectral information is directly mapped onto the horizontal axis of the DMD, we can acquire an arbitrary number spectral bands by modulating selective DMD pixel columns and choose a desired spectral resolution by binning the DMD pixels. Thereby, our method allows for both multi-spectral imaging (discrete or overlapped spectral bands) and hyperspectral imaging (narrow spectral bands over a continuous spectral range) on demand without requiring any hardware modification.

Single-pixel spectral time-of-flight imaging (x, y, λ, t)

To measure high-D light field information, conventional multidimensional optical imaging methods often trade in light information along another axis. In contrast, in a cascaded compressed-sensing camera, we employ an optical architecture that first reduces the spatial dimensionality of a light datacube, followed by transforming high-D light information to refill this spatial axis, thereby enabling a complete measurement of optical properties. Here, we demonstrate a four-dimensional (4D) (x, y, λ, t) spectral time-of-flight imaging by fully exploiting the bandwidth advantage of single-pixel detectors. We imaged a scene of colored objects positioned at different depths (Fig. 4a). The distance between the objects and the imaging system varied from 0.6 to 0.9 m.

To cover a broad spectrum in the visible light range (500–650 nm), we illuminated the objects with a tunable pulsed laser (a femtosecond laser equipped with an optical parametric amplifier) stepwise, recorded the signals in each spectral band, and summed the raw data for processing. The central wavelengths of the illuminating beams were set from 500 nm to 650 nm with 10 nm increments, resulting in a total of 16 illuminations. The spectral bandwidth of each laser pulse is approximately 20 nm, providing a sufficient

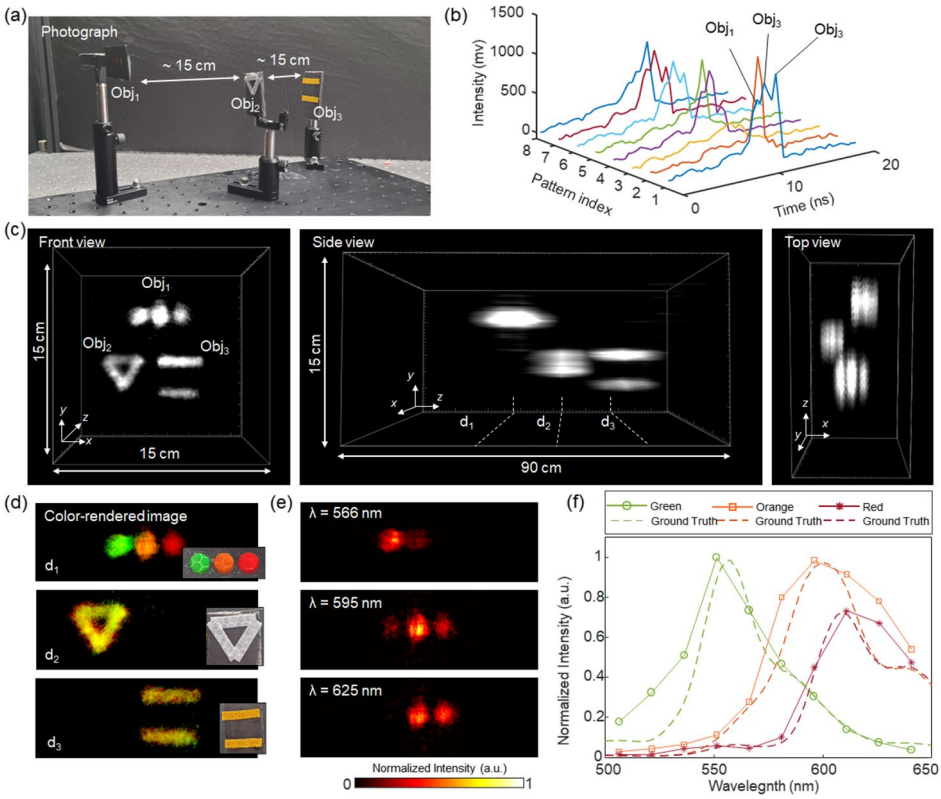

Fig. 4 Single-pixel full-waveform spectral time-of-flight imaging. **a** Photogram of the scene. The scene consists of three objects at different depths. **b** Representative temporal waveforms. **c** Reconstructed 3D scene visualized by the maximum intensity projection. **d** Representative depth images. The images are color-rendered by using their spectral irradiance ranging from 500 nm to 650 nm. The insets show the ground truth photographs at the corresponding depths. **e** Spectral images at depth d_1 at representative wavelengths. **f** Spectra of the colored circles at depth d_1

spectral overlap between adjacent illuminations. The back-scattered light from the objects was captured by a two-stage compressed-sensing single-pixel camera (Fig. 3a) equipped with a high-speed APD (1.6 GHz bandwidth) and a digitizer (2.5 GHz digital sampling rate). The DMD was synchronized with the mode-lock frequency (1 kHz) of the femtosecond laser. The temporal waveforms of the raw data from the 16 measurements were integrated without any post-processing. We used a total of 45 image rotation angles and 192 Hadamard patterns, enabling the recording of 86,400 time-of-flight waveforms with a spatial SR of ~ 0.075 (Fig. 4b).

A 4D (x, y, λ, t) datacube with $(384 \times 384 \times 10 \times 15)$ voxels was reconstructed from the time-of-flight waveforms obtained and visualized in Fig. 4c. The temporal sampling rate of the digitizer is 2.5 GHz, leading to an axial sampling rate of ~ 6 cm ($c\Delta t/2$; c, speed of light in the air; Δt, time of flight), considering the round-trip travel time of the illuminating pulse. The 3D rendering recovers both the shape and depth of the object with high fidelity (Fig. 4c). The lateral field of view depends on the distance between the camera and the object because of a varying magnification, and it approximates 15 cm × 15 cm in the middle. The spectrally and temporally integrated 2D scene matches well with the reference image captured by a color camera (Fig. 4d). Figure 4e shows representative

images at different wavelengths. The RMSE between the normalized spectrum and the ground truth of the green, orange, and red objects are 0.12, 0.14, 0.07, respectively.

Parallel acquisition

Although the proposed system efficiently captures a multidimensional light datacube with a high compression ratio, the imaging speed is limited by the refresh rate of the spatial light modulator and the rotating speed of the Dove prism. In our demonstration, the refresh rate of the DMD was limited to 8 kHz, creating a bottleneck in the imaging speed as several tens of thousands of patterns were displayed by the DMD per measurement (Table S1). This limitation could be mitigated by using other advanced high-speed spatial light modulators such as high-speed DMDs [18], light vavles [44] and metasurfaces [45]. However, when increasing the refresh rate of the spatial light modulator, the high moment of the inertia of the rotating Dove prism may pose a new challenge, impeding fast acquisition.

To overcome this limitation, we demonstrate proof-of-concept experiments on cascaded compression without mechanical scanning (Fig. 5). Instead of using a rotating Dove prism, we devised a light field tomography module (Fig. 5b) to map the *en-face* projections of the light field onto the DMD in a parallel manner (see Materials and Methods). The module is consisting of 12 Dove prisms with varied orientations and 4 cylindrical lenses. The Dove prism array placed at the back focal plane of an imaging objective lens captures 12 perspective views of the same scene. The relaying cylindrical lenses integrate the perspective images along the horizontal axis. The DMD sequentially encodes the image with vertical Hadamard patterns and the signal is then integrated by the relaying optics and the single-pixel detector. In this configuration, the acquisition

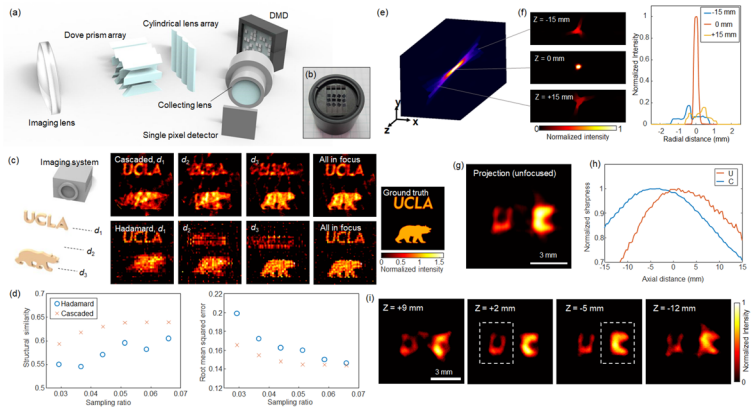

Fig. 5 4D (x, y, u, v) light field imaging using a single-pixel camera. **a** The light field tomography module, comprising 4 cylindrical lenses and a 3 × 4 array of Dove prisms, captures a 4D light field using a single-pixel camera. The incident light is divided into 12 perspective views, and its *en-face* parallel projections are mapped onto the DMD. **b** Photograph of the light field tomography module showing the 2 mm × 2 mm Dove prisms and the 15 mm focal length of the cylindrical lenses. **c** Numerical simulation comparing light field imaging using the cascaded compression method and conventional single-pixel imaging with a Hadamard basis. The enlarged view is provided in Supplementary Fig. S4. **d** RMSE and SSIM of the reconstructed scene. **e** A 4D light field of an optical focus captured by the single-pixel camera, with the 3D scene visualized using maximal intensity projection. **f** The horizontal cross-section of the image at different depths presented as a graph. **g** A physical target displaying the alphabet 'U' and 'C' imaged by the single-pixel camera. The distance between the alphabets is approximately 7 mm. The axially integrated image shows blurred edges of the alphabet. **h** Sharpness analysis of the image along the depth. **i** Computationally refocused images at different depths, with the images at maximum sharpness depths highlighted

speed is solely dictated by the refresh rate of the DMD. Moreover, the parallel array of Dove prisms enables 3D light-field imaging.

We first conducted a numerical simulation to compare the cascaded compression method and conventional compressed single-pixel light-field imaging employing a Hadamard basis in terms of image reconstruction fidelity (Fig. 5c, d and Supplementary Fig. S4). The cascaded compression method outperforms the Hadamard basis method when the compression ratio is high (SR $\leq$ 0.07). Next, we experimentally demonstrated single-pixel light-field imaging. We choose a total sampling rate of 3072 (12 Dove prism angles and 256 Hadamard patterns), which is corresponding to the SR of 0.06. We captured an optical focus (Fig. 5e, f) and visualize it in 3D space by a maximal intensity projection method. We also captured the scene generated by a 3D-printed physical target. (Fig. 5g-i). The target is consisting of two optical masks displaying alphabet 'U' and 'C' where their distance is approximately 7 mm. Based on the image sharpness analysis, we can find the depth location of the optical masks (Fig. 5h). With the light field tomography module, we successfully demonstrated cascaded compressed single-pixel imaging without rotating Dove prisms.

Discussion

The architecture of cascaded compressed imaging is scalable, allowing high-D optical information to be mapped to a low-D space stagewise. Figure 6 shows a generalized scheme with multiple compression stages, each consisting of a low-D folding and a high-D unfolding step. In the low-D folding step, the image is spatially compressed along the x-axis through optical transformation, creating a blank low-D axis. Next, in the unfolding step, we map the optical information along a high-D axis to this blank axis and measure the resultant image using computed tomography. Repeating these two steps at each compression stage allows measurement of an additional photon characteristic, and cascading multiple stages leads to a complete characterization. In the final stage, we further reduce the data dimensionality to 0D by modulating the image using a spatial light modulator and acquiring the spatially integrated signals using a single-pixel detector. Although we demonstrated only a two-stage system for (x, y, λ) and (x, y, λ, t) imaging, the cascaded compression framework can be expanded to measuring other photon characteristics, such as polarization or depth, by using a Wollaston prism [46] or a volume hologram [47], respectively, as a mapping device during high-D unfolding.

Cascaded compressed imaging is highly efficient in measuring a high-D light datacube. For conventional multidimensional imaging, acquiring optical information along an extra dimension often yields a significant dataload, posing challenges in data transfer, storage, and processing. The cascaded compression alleviates this problem by mapping the optical information along the axis of interest to a low-D spatial axis in a compressed manner rather than measuring the complete information. Such a measurement scheme is particularly suitable for multidimensional imaging because it exploits the sparsity of natural scenes in a high-D space that cannot be commonly found in a low-D space. For example, in 4D full-waveform hyperspectral time-of-flight imaging, the spectrally resolved image at each time instant of the reflected time-of-flight signal is far simpler than a conventional color photograph, where the light signals are temporally and spectrally integrated in the same image. As a result,

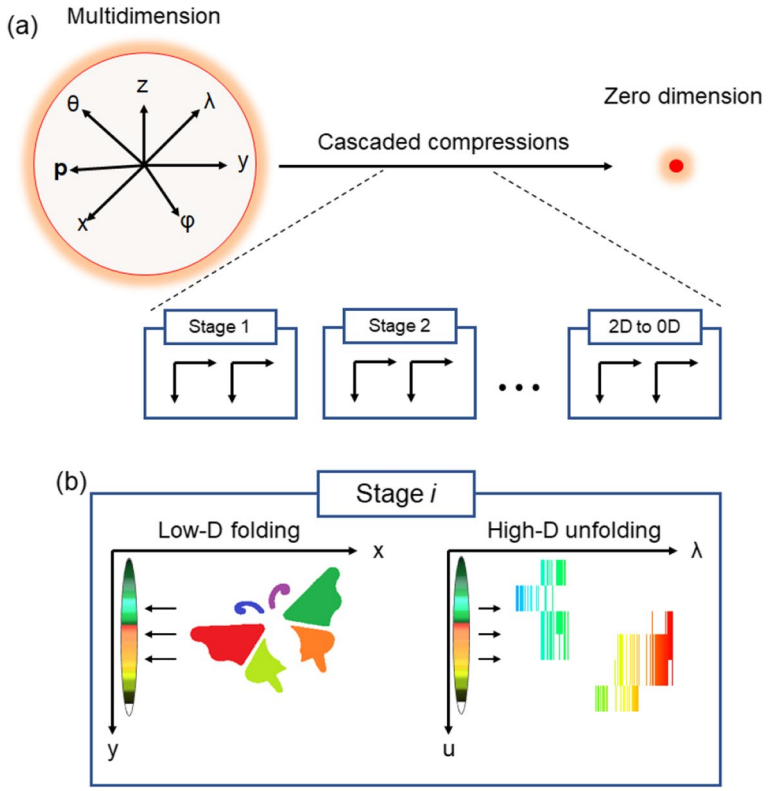

Fig. 6 Conceptual framework of cascaded compressed imaging. **a** High-D optical information is mapped into a low-D space by a series of optical operations. **B** Low-D folding and high-D unfolding in a measurement stage. In this example, a 2D image is first transformed into a 1D line during the low-D folding process, creating a blank spatial axis. The spectral information is then mapped to this axis during the high-D unfolding process

the optical information along the unfolded spectral and temporal axes is intrinsically sparse, and it can be measured at a much lower rate than the Nyquist sampling condition. Given this prior knowledge on signal sparsity, we can faithfully recover a high-D datacube with only a small fraction of information recorded. The same principle is also applicable to acquiring other light datacube dimensions. For example, we demonstrated 4D(x, y, u, v) compressed light field imaging and presented the results in Fig. 5.

Lastly, the competitive advantage of cascaded compressed imaging over conventional DMD-based single-pixel cameras exists not only in high-D imaging but also in low-D imaging. For example, when capturing a 2D image with conventional Hadamard encoding, the limited onboard memory size of DMDs usually leads to insufficient samplings for high-resolution imaging. As a result, the image size of most single-pixel cameras using a DMD is limited to a few tens of thousands of pixels, far fewer than the resolution of a conventional digital camera. In contrast, by decomposing the measurement into two stages, we loaded only 1,920 patterns on the DMD and captured 172,800 samples. Conventional DMD-based single-pixel cameras cannot achieve this because the number of samples acquired exceeds the modulation capability of the DMD (64 Gbit, maximum 55,924 patterns).

In conclusion, we developed a new category of compressed-sensing cameras capable of capturing a multidimensional optical field using a single-pixel detector. By mapping high-D information to a low-D space through a cascaded process, we enable a highly efficient measurement of a light datacube with rich information contents. We expect our method will find broad multidimensional imaging applications that use single-pixel detectors.

Materials and methods

Experimental setup

We collected the image using a Dove prism (PS992M, Thorlabs) and cylindrical lens ($f = 100$ mm) with an optical power along the vertical direction (Fig. 3a). The Dove prism was mounted on a high-speed rotary stage (ACR32UT, IntelLiDrives). The image was then projected onto a DMD (V-650 L, Vialux) through a 4-f telescopic imaging system with a unity magnification ($f = 75$ mm). In practice, the DMD was mounted on a stage inclined at 45 degrees so that the reflection beam direction is parallel to the optical table. A diffraction grating (GT-25-03, 300 grooves/mm) was positioned at the Fourier plane of the imaging system to disperse the light along the horizontal direction. We can encode the image with a desired binary pattern by modulating the micromirrors on the DMD.

The light-collecting optics after the DMD consist of a 4-f telescoping imaging system with a diffraction grating at the Fourier plane, which un-disperse the beam. The focal length of the lenses and the groove density of the diffraction grating are the same but with a reversed dispersion direction. After passing the light-collecting optics, the light is directed to the single-pixel detector (APD210, 1.6 GHz, MeloSystems) and digitized by the high-speed digitizer (6426E, 2.5 GHz, Picoscope) for processing.

For time-of-flight imaging, we use a Ti: Sapphire femtosecond laser (Astrella-F-1 K, 808 nm, 100 fs pulse width, Coherent) as an illumination source. An optical parametric amplifier (TOPAS-Prime, Light Conversion) further converts the laser wavelength to the visible light range. The trigger-out from the mode-locked laser was used as a reference for accurate timing for the digitizer.

System calibration

A lateral shift of the image at the DMD plane often accompanies the rotation of the Dove prism in our system due to misalignment of the system's optical axis with respect to the center of rotation of the Dove prism. The deviation was corrected computationally by introducing a lateral shifting operator into the linear forward model.

To determine the amount of the lateral shift of the image varied with the Dove prism rotation angle, we positioned a pinhole at the optic axis of the system, illuminated it with a narrow band laser beam ($\lambda = 532$ nm), and captured the image using a conventional Hadamard basis single-pixel imaging configuration. The exact spatial location of the mapped laser beam was recorded, and the deviation from the center of the DMD was calculated at different rotation angles of the Dove prism. The corresponding lateral shifting operator was compensated when reconstructing the image.

The spectral radiance response was calibrated using a white lamp and a benchmark spectrometer (STS-VIS-L-25-400-SMA, Ocean Optics). An iris was positioned at the object plane of the system and illuminated by a white lamp. The light was dispersed

along the horizontal direction and mapped onto the DMD. The DMD was divided into 90 vertical stripes with a horizontal pixel size of 4 pixels. We sequentially turned on the vertical stripe images on the DMD and recorded the intensity value by the single-pixel detector for each stripe. The spectral response of the system was normalized by comparing it with the spectrum measured by a benchmark spectrometer.

Light field tomography module

A schematic of a light field tomography module is illustrated in Fig. 5a. The module consists of 12 Dove prisms (height = 2 mm, length = 8.4 mm, arranged in 3 rows and 4 columns with 3 mm spacing between adjacent prisms) with different orientations ($\theta = 0°$, 7.5°, 15°, 22.5°,· · ·,82.5°). The Dove prism array is located at the back focal plane of the imaging objective lens. Each Dove prism captures a different region of Fourier spectrum of the scene, or equivalently, each prism observes the same scene from a different perspective. As the dove prism rotates the input image by an angle of 2θ, the resulting 12 images cover angles ranging from 0° to 180° spaced equally. The images then pass through four cylindrical lenses (height = 15 mm, width = 2.8 mm, focal length along the horizontal axis = 15 mm) and mapped onto the DMD. The resultant images are 12 perspective images that are rotated and spatially integrated along the vertical axis. The DMD sequentially encodes the images with a Hadamard pattern and relays the light to the single-pixel detector.

Digital refocusing can be done by shifting the captured perspective views according to the desired focal plane and the dove prism location during the image reconstruction process [33]. Specifically, to reconstruct a defocused image at depth Δf from the focal plane, the Hadamard pattern $\mathbf{V}_j$ is vertically translated by $1/p \cdot \Delta f/(f + \Delta f) \cdot (u_x \sin\theta - u_y \cos\theta)$, where p is the DMD pixel pitch and f is the focal length of the imaging objective lens. u_x and u_y are spatial location of the dove prism along the x and y direction with respect to the center of the imaging objective lens. θ is the rotating angle of the scene (half the rotated angle of the dove prism).

Image reconstruction

To reconstruct the image from the measurement, we solved the linear inverse problem of Eq. (2). We adopted a fast iterative shrinkage-thresholding algorithm [41] with a proximal operator of the l_1-norm [35] for the regularization, which lessens the computational cost compared to the use of state-of-art regularizers such as total variation and wavelet transformations. The hyperparameter μ was chosen between 0.05 and 0.5, which does not significantly affect the image reconstruction fidelity. The reconstruction time for high-dimensional scene of size ($384 \times 384 \times 10$) is approximately 4.7 s for 30 iterations using a personal computer (Intel i9-10900 K CPU, NVIDIA GeForce RTX 3070 GPU). The simulation code is provided in Supplementary Information.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43074-024-00152-5.

Supplementary Material 1.

Acknowledgements
Not applicable.

Authors' contributions
J.P. and L.G. designed the system. J.P. built the system, performed the experiments and analysed the data. J.P. and L.G. wrote the manuscript. L.G. supervised the project.

Funding
This work was supported partially by National Institutes of Health (R35GM128761).

Data availability
The algorithms and simulation codes for this study are available at Supplementary Information. The specific implementation of the codes and data that supports this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 24 June 2024 Revised: 16 October 2024 Accepted: 25 October 2024
Published online: 07 November 2024

References

1. Edgar MP, Gibson GM, Padgett MJ. Principles and prospects for single-pixel imaging. Nat Photonics. 2019;13:13–20.
2. Gibson GM, Johnson SD, Padgett MJ. Single-pixel imaging 12 years on: a review. Opt Express. 2020;28:28190–208.
3. Wu G, et al. Light field image processing: an overview. IEEE J Sel Top Signal Process. 2017;11:926–54.
4. Savage N. Digital spatial light modulators. Nat Photonics. 2009;3:170–2.
5. Shapiro JH. Computational ghost imaging. Phys Rev A. 2008;78:061802.
6. Duarte MF, et al. Single-pixel imaging via compressive sampling. IEEE Signal Process Mag. 2008;25:83–91.
7. Chan WL, et al. A single-pixel terahertz imaging system based on compressed sensing. Appl Phys Lett. 2008;93:121105.
8. Studer V, et al. Compressive fluorescence microscopy for biological and hyperspectral imaging. Proc Natl Acad Sci. 2012;109:E1679–87.
9. Zhang Z, Ma X, Zhong J. Single-pixel imaging by means of Fourier spectrum acquisition. Nat Commun. 2015;6:1–6.
10. Edgar M, et al. Simultaneous real-time visible and infrared video with single-pixel detectors. Sci Rep. 2015;5:1–8.
11. Sun M-J, et al. Single-pixel three-dimensional imaging with time-based depth resolution. Nat Commun. 2016;7:1–6.
12. Yu H, et al. Fourier-transform ghost imaging with hard X rays. Phys Rev Lett. 2016;117:113901.
13. Stockton P, et al. Tomographic single pixel spatial frequency projection imaging. Opt Commun. 2022;520;021907.
14. Candes EJ, Tao T. Near-optimal signal recovery from random projections: Universal encoding strategies? IEEE Trans Inf Theory. 2006;52:5406–25.
15. Torabzadeh M, Park I-Y, Bartels RA, Durkin AJ, Tromberg BJ. Compressed single pixel imaging in the spatial frequency domain. J Biomed Opt. 2017;22:030501.
16. Horadam KJ. Hadamard matrices and their applications. Hadamard matrices and their applications. New Jersey: Princeton University Press; 2012.
17. Zhang Z, Wang X, Zheng G, Zhong J. Hadamard single-pixel imaging versus Fourier single-pixel imaging. Opt Express. 2017;25:19619–39.
18. Dudley D, Duncan WM, Slaughter J. Emerging digital micromirror device (DMD) applications. MOEMS Display Imaging Syst. 2003;4985:14–25 (International Society for Optics and Photonics).
19. Sun MJ, Meng LT, Edgar MP, Padgett MJ, Radwell N. A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging. Sci Rep. 2017;7:1–7.
20. López-García L, et al. Efficient ordering of the Hadamard basis for single pixel imaging. Opt Express. 2022;30:13714–32.
21. Duarte MF, Eldar YC. Structured compressed sensing: from theory to applications. IEEE Trans Signal Process. 2011;59:4053–85.
22. Dong W, Shi G, Li X, Ma Y, Huang F. Compressive sensing via nonlocal low-rank regularization. IEEE Trans Image Process. 2014;23:3618–32.
23. Yu X, Stantchev RI, Yang F, Pickwell-MacPherson E. Super sub-nyquist single-pixel imaging by total variation ascending ordering of the hadamard basis. Sci Rep. 2020;10:1–11.
24. Gao L, Smith RT. Optical hyperspectral imaging in microscopy and spectroscopy–a review of data acquisition. J Biophotonics. 2015;8:441–56.
25. Park J, Feng X, Liang R, Gao L. Snapshot multidimensional photography through active optical mapping. Nat Commun. 2020;11:1–13.
26. Liang J, Wang P, Zhu L, Wang LV. Single-shot stereo-polarimetric compressed ultrafast photography for light-speed observation of high-dimensional optical transients with picosecond resolution. Nat Commun. 2020;11:5252.
27. Zhao Z, et al. Redundant compressed single-pixel hyperspectral imaging system. Opt Commun. 2023;546:129797.

28. Jin S, et al. Hyperspectral imaging using the single-pixel fourier transform technique. Sci Rep. 2017;7:1–7.
29. Bian L, et al. Multispectral imaging using a single bucket detector. Sci Rep. 2016;6:1–7.
30. Sun B, et al. 3D computational imaging with single-pixel detectors. Science. 2013;340:844–7.
31. Helgason S, Helgason S. The radon transform, vol. 2. New York: Springer; 1980.
32. Feng X, Gao L. Ultrafast light field tomography for snapshot transient and non-line-of-sight imaging. Nat Commun. 2021;12:2179.
33. Ma Y, et al. Light-field tomographic fluorescence lifetime imaging microscopy. Proc Natl Acad Sci. 2024;121:e2402556121.
34. Li X, Luo S. A compressed sensing-based iterative algorithm for CT reconstruction and its possible application to phase contrast imaging. Biomed Eng OnLine. 2011;10:73.
35. Kudo H, Suzuki T, Rashed EA. Image reconstruction for sparse-view CT and interior CT—introduction to compressed sensing and differentiated backprojection. Quant Imaging Med Surg. 2013;3:147–61.
36. Candes E, Romberg J. Sparsity and incoherence in compressive sampling. Inverse Probl. 2007;23:969.
37. Xu J, Pi Y, Cao Z. Optimized projection matrix for compressive sensing. EURASIP J Adv Signal Process. 2010;2010:1–8.
38. Abo-Zahhad MM, Hussein AI, Mohamed AM. Compressive sensing algorithms for signal processing applications: a survey. Int J Commun Netw Syst Sci. 2015;8:197.
39. Arjoune Y, Kaabouch N, Ghazi E, Tamtaoui A. A performance comparison of measurement matrices in compressive sensing. Int J Commun Syst. 2018;31:e3576.
40. Nouasria H, Et-tolba M, Bedoui A. New sensing matrices based On orthogonal hadamard matrices for compressive sensing. in 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE; 2019. p. 186–91.
41. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci. 2009;2:183–202.
42. Higham CF, Murray-Smith R, Padgett MJ, Edgar MP. Deep learning for real-time single-pixel video. Sci Rep. 2018;8:1–9.
43. Li Z, et al. Efficient single-pixel multispectral imaging via non-mechanical spatio-spectral modulation. Sci Rep. 2017;7:41435.
44. Tzang O, et al. Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform. Nat Photonics. 2019;13:788–93.
45. Shaltout AM, Shalaev VM, Brongersma ML. Spatiotemporal light control with active metasurfaces. Science. 2019;364:eaat300.
46. Oliva E. Wedged double Wollaston, a device for single shot polarimetric measurements. Astron Astrophys Suppl Ser. 1997;123:589–92.
47. Luo Y, et al. Laser-induced fluorescence imaging of subsurface tissue structures with a volume holographic spatial-spectral imaging system. Opt Lett. 2008;33:2098–100.

Publisher's note