Tutorial on compressed ultrafast photography

Yingming Lai[iD],[†] Miguel Marquez[iD],[†] and Jinyang Liang[iD]*
Université du Québec, Institut National de la Recherche Scientifique, Centre Énergie Matériaux
Télécommunications, Laboratory of Applied Computational Imaging, Varennes, Québec, Canada

ABSTRACT. **Significance:** Compressed ultrafast photography (CUP) is currently the world's fastest single-shot imaging technique. Through the integration of compressed sensing and streak imaging, CUP can capture a transient event in a single camera exposure with imaging speeds from thousands to trillions of frames per second, at micrometer-level spatial resolutions, and in broad sensing spectral ranges.

Aim: This tutorial aims to provide a comprehensive review of CUP in its fundamental methods, system implementations, biomedical applications, and prospect.

Approach: A step-by-step guideline to CUP's forward model and representative image reconstruction algorithms is presented with sample codes and illustrations in Matlab and Python. Then, CUP's hardware implementation is described with a focus on the representative techniques, advantages, and limitations of the three key components—the spatial encoder, the temporal shearing unit, and the two-dimensional sensor. Furthermore, four representative biomedical applications enabled by CUP are discussed, followed by the prospect of CUP's technical advancement.

Conclusions: CUP has emerged as a state-of-the-art ultrafast imaging technology. Its advanced imaging ability and versatility contribute to unprecedented observations and new applications in biomedicine. CUP holds great promise in improving technical specifications and facilitating the investigation of biomedical processes.

1 Introduction

Optical imaging of transient events in their actual time of occurrence exerts compelling scientific significance and practical merits.[1] Occurring in two-dimensional (2D) space and at femtosecond (1 fs = 10^{-15} s) to microsecond (1 μs = 10^{-6} s) time scales, these transient events reflect many important fundamental mechanisms in biology.[2–4] However, many transient phenomena are either nonrepeatable or difficult to reproduce. Examples include the spontaneous synaptic activities,[5] nanoparticles' luminescence lifetime at different temperatures,[6] and light scattering in living tissue.[7] Under these circumstances, the conventional pump–probe methods, requiring numerous repeatable experiments, are inapplicable. Meanwhile, the pump–probe approaches sense photons' time-of-arrival using complex apparatus to perform time-consuming scanning in either space or time. In these cases, even if the transient phenomena are reproducible, these

---

*Address all correspondence to Jinyang Liang, jinyang.liang@inrs.ca

[†]These authors contributed equally to this work.

methods would suffer from substantial inaccuracy due to experimental perturbation and low productivity due to the events' low occurrence rates.

Single-shot ultrafast optical imaging techniques[8,9] can overcome these limitations by capturing the entire dynamic process in real time (i.e., in the actual duration of the event's occurrence) without repeating measurements. Benefiting from advancements in optoelectronics, laser science, information theory, and computational techniques, single-shot ultrafast optical imaging has become a burgeoning research field in the past decade. Thus far, the mainstream techniques can be generally categorized into the domains of active illumination and passive detection. For the former, temporal information of a dynamic scene is mapped into an optical marker (e.g., spectrum and spatial frequency) of one or multiple ultrashort probe pulses. On the detection side, appropriate devices and methods (e.g., color filter and spatial Fourier transformation) are used to extract the corresponding optical marker, which deduces the scene's evolution. These active-illumination-based approaches feature femtosecond temporal resolution by leveraging ultrashort durations of ultrafast probe pulses and provide high sensitivity by being compatible with advanced cameras based on the charge-coupled device (CCD) or complementary metal–oxide semiconductor (CMOS) technology. Nonetheless, they cannot capture the self-luminescence scenes, including dynamic scattering,[10] photoluminescence intensity decay,[11] and plasma emission.[12] Passive detection can overcome this limitation. In this category, receive-only ultrafast detectors are used to record the emitted and/or scattered photons from the dynamic scene. Various mechanisms, including Kerr-effect-based time gating,[13] deflection of moving photoelectrons by a varying electrical field,[14] and charge transfer in a series of registers,[15] have been used to provide ultrahigh temporal resolution. Meanwhile, the inferior bandwidth of electronics to its optical counterpart caps the ultimate imaging speed of these passive-detection approaches lower to the active-illumination modalities. Overall, the active illumination and passive detection approaches often carry highly complementary technical specifications. Altogether, they incessantly expand the human vision to see previous inaccessible events.

Among existing techniques, compressed ultrafast photography (CUP) has emerged as a potent single-shot ultrafast optical imaging modality.[16] Invented in 2014 in Dr. Lihong V. Wang's laboratory,[17] CUP innovatively synergizes compressed sensing (CS) with streak imaging. Leveraging the sparsity existing in the targeted scenes, the operation of this hybrid approach includes physical data acquisition followed by computational image reconstruction.[18,19] In data acquisition, the light from a 2D dynamic scene is recorded in one or more snapshots in a single shot via a CS paradigm containing spatial encoding, temporal shearing, and spatiotemporal integration. Different from conventional ultrafast imaging, the acquired snapshot often bears no resemblance to the scene. Then, the snapshot is input into an algorithm to retrieve the movie of the target dynamic scene by solving a minimization problem.[20]

CUP provides many attractive conceptual novelties and practical advantages. First, the spatial encoding and temporal shearing operations allow a mixture of information between time and space, which enables CUP to have a large sequence depth (i.e., the number of frames in each recorded movie) compared with other single-shot ultrafast imaging systems based on spatial frequency multiplexing,[21–24] spectral filtering,[25–30] and beam splitting.[31–34] Meanwhile, it overcomes the limitations in sensing dimension in conventionally regarded one-dimensional (1D) high-speed sensors.[14,35] Compared with ultrafast CCD sensors that have a low fill factor, CUP uses spatiotemporal multiplexing to effectively enhance the light throughput in data acquisition, which improves the feasibility of image reconstruction.[36] It is compatible with many scientific-grade CCD/CMOS sensors without interrupting their normal operations, which retains their responsive spectrum and sensitivity while endowing them with ultrahigh speeds.[37] Second, its generic sensing paradigm can be embodied in both active-illumination and passive-detection schemes. Each major operation (i.e., spatial encoding, temporal shearing, and spatiotemporal integration) can be optically realized by various devices, indicating high design flexibility, multi-spatiotemporal-scale imaging ability, and broad spectral coverage. Third, computational image reconstruction, as an indispensable step in CUP's operation, lifts certain burdens in system design from hardware. Advances in CS,[38] machine learning,[39] and information theory[40] can be directly implemented in CUP's image reconstruction. Finally, CUP exhibits light throughput advantages by capturing information in two spatial dimensions and time simultaneously in a single exposure. In contrast, the multiple-shot methods can only collect information

from a column (from point scanning) or a slice (from line scanning) of the datacube.[41] Meanwhile, distinguished from single-shot framing (or mapping) photography,[30,34,42] CUP maintains time continuity in data acquisition, which further enhances the amount of acquirable information.[18]

Because of its unprecedented imaging ability, CUP immediately became a research focus since its invention. New designs in hardware and innovative development of image reconstruction are being reported frequently. New applications in biomedicine, physics, and engineering are highlighted. Comprehensive reviews of CUP can be found in the literature.[16,43] Other reviews of CUP are included in the surveys of ultrafast imaging technologies.[1,8,9,18,44–48] However, thus far, there has not been a practical guide for developing CUP systems using an anatomy fashion. Thus, in this tutorial, we first review the operating principle of CUP with simulation examples (in Matlab and Python) to guide readers on how to generate compressively recorded snapshots from a spatiotemporal datacube using the forward model as well as how to reconstruct the spatiotemporal datacube from the snapshots using representative methods in analytical-modeling-based approach and machine learning. Then, we will provide an extensive survey of existing methods for each of the major operations in CUP's sensing paradigm—spatial encoding, temporal shearing, and spatiotemporal integration. Afterward, we will discuss the representative applications of CUP in biomedicine. Finally, we summarize CUP's accomplishments and provide the prospect of its future development.

2 Method

A schematic of dual-view CUP is shown in Fig. 1. In data acquisition, a dynamic scene is imaged by front optics and split into two arms. The transmitted component forms the image of the dynamic scene on a spatial encoder. Unlike many other compressive temporal imaging modalities that use multiple fast-changing patterns during image acquisition,[49–54] a single static pattern is used for CUP's spatial encoder. Then, the frames in the spatially encoded scene are deflected by a temporal shearing unit to different spatial positions along the sweeping direction. Finally, the encoded and sheared scene is spatiotemporally integrated by a sensor, producing a compressive 2D snapshot, which is defined as the time-sheared view and used hereafter in this tutorial. This paradigm to capture the time-sheared view was implemented in the original CUP configuration.[17] In the ensuing implementations, it was found that a direct capture of a time-integrated snapshot of the dynamic scene could enhance the reconstructed imaging quality.[55] Defined as the time-unsheared view and used hereafter in this tutorial, this snapshot outlines the region of occurrence of the dynamic scene, which reduces the number of unknowns for image reconstruction and facilitates its convergence to the optimal result. It is particularly useful when the dynamic scene occurs on a static or slowly moving object (e.g., intensity decay of photoluminescence emitted from nanoparticle-labeled cells[56]). It is noted that CUP systems of more than two views have been featured in recent progress to further boost image quality.[57–66] For example, lossless-encoding CUP contains the time-unsheared view and two complementary time-sheared views.[58] Nonetheless, the formation of these views shares similar data acquisition paradigms as the ones described above and thus is not discussed here.

To assist readers in comprehending CUP's paradigm, in Secs. 2.1 and 2.2, we provide in-depth theoretical derivation and simulation. The presented examples are meticulously designed
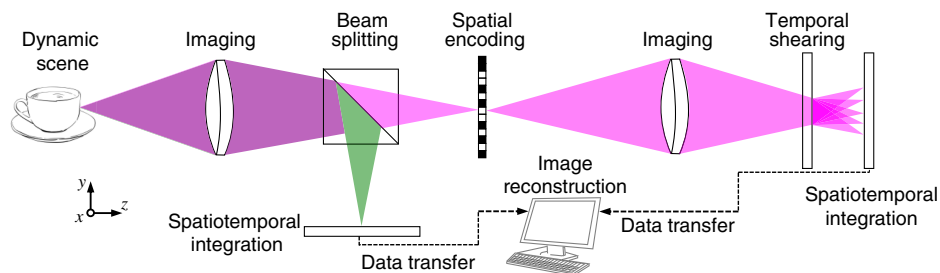

Fig. 1 Operating principle of dual-view CUP. The illustration depicts the beam paths for time-sheared and time-unsheared views, represented by magenta and green colors, respectively.

to use basic features and functions in Matlab (R2020b) and Python (version 3.9). Two versions of Python codes are prepared for readers with different levels of programming experience.

2.1 Forward Model

The forward model of CUP formulates the process of recording a three-dimensional [3D; i.e., (x, y, t)] scene to one or a few 2D snapshots. In general, this forward model can be expressed mathematically using either element-wise or matrix-vector notations.

2.1.1 *Element-wise notation*

Many mathematical and scientific libraries are designed to efficiently handle element-wise operations. The dynamic scene and the binary-valued encoding mask are denoted by $\mathbf{F} \in \mathbb{R}^{M \times N \times L}$ and $\mathbf{R} \in \mathbb{R}^{M \times N}$, respectively. M and N represent the data lengths in the two spatial dimensions, and L signifies the data length in time. The discrete output from the sensor for the time-sheared view (hence the subscript "ts") can be modeled as

$$(G_{\text{ts}})_{i,j} = \sum_{l=0}^{L-1} \overline{F}_{i,j,l} \overline{R}_{i,j} + (E_{\text{ts}})_{i,j}, \tag{1}$$

where

$$\overline{F}_{i,j,l} = \begin{cases} F_{i,j-l,l} & \text{if } l \leq j \leq [N + (l-1)] \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\overline{R}_{i,j} = \begin{cases} R_{i,j-l} & \text{if } l \leq j \leq [N + (l-1)] \\ 0 & \text{otherwise,} \end{cases}$$

where $\overline{F}_{i,j,l}$ is the intensity of the (i, j, l)'th element of a right-zero-padded version with a frame-dependent right circular shifting of the dynamic scene's datacube with $\overline{\mathbf{F}} \in \mathbb{R}^{M \times [N+(L-1)] \times L}$. $\overline{R}_{i,j}$ stands for the intensity of the (i, j)'th element a right-zero-padded version with a frame-dependent right circular shifting of the spatial encoder with $\overline{\mathbf{R}} \in \mathbb{R}^{M \times [N+(L-1)]}$. $(G_{\text{ts}})_{i,j}$ is the intensity measured on the (i, j)'th element of the sensor with $\mathbf{G}_{\text{ts}} \in \mathbb{R}^{M \times [N+(L-1)]}$. $(E_{\text{ts}})_{i,j}$ stands for the noise of the (i, j)'th element in $\mathbf{G}_{\text{ts}}$ with $\mathbf{E}_{\text{ts}} \in \mathbb{R}^{M \times [N+(L-1)]}$.

The discrete output for the time-unsheared view (hence "tu" as the subscript) can be modeled as

$$(G_{\text{tu}})_{i,j} = \sum_{l=0}^{L-1} F_{i,j,l} + (E_{\text{tu}})_{i,j}, \tag{2}$$

where $(G_{\text{tu}})_{i,j}$ is the intensity measured at the (i, j)'th element of the time-unsheared view with $\mathbf{G}_{\text{tu}} \in \mathbb{R}^{M \times N}$, and $(E_{\text{tu}})_{i,j}$ represents the noise in $\mathbf{G}_{\text{tu}}$ with $\mathbf{E}_{\text{tu}} \in \mathbb{R}^{M \times N}$.

As an example, a Matlab script that simulates dual-view CUP's forward model, with a linear shearing operator and a pseudorandom binary mask, is shown in Algorithm 1. Moreover, a step-by-step guide with illustrations of the "cell-division" dynamic scene is shown in Fig. 2. The ground truth video was taken from the public "Mouse Embryo Tracking" database[67] and can be downloaded using the link in Ref. 68.

2.1.2 *Matrix-vector notation*

The element-wise notation of CUP's forward model, despite owning simple expression and easy comprehension, is inherently limited by its sequential execution. This characteristic engenders surplus calculations within specific functions when the modeling is subjected to extensive datasets or algorithms mandating intricate computations—such as matrix inversion, matrix factorization, eigenvalue decomposition, and low-rank approximation. Thus, most practices use matrix-vector operations by converting $\mathbf{F}$ into vector $\mathbf{f} \in \mathbb{R}^{n \times 1}$, where $n = M \cdot N \cdot L$. In this

Algorithm 1 Simulating dual-view CUP's forward model with the element-wise notation using Matlab.

```
%% Example of dual-view CUP's sensing process

% Encoding step (generating R)

load('Cell.mat')                           % Loading the example video – F

[M,N,L] = size(F);                         % Calculating the video dimensions
                                           %  (y,x,t)->(M,N,L)

R = 1*(rand(M,N)>0.5);                     % Mask initialization with a
                                           %  transmittance of ~50%

R = repmat(R,1,1,L);

Gts = F.* R;                               % Spatial encoding – Hadamard product

Gts = padarray(Gts,[0,L-1,0],0,'post');    % Right column zero padding

for l=0:L-1                                % Shearing operation

  Gts(:,:,l+1) =circshift(Gts(:,:,l+1),[0,l]);

end

Gts =sum(Gts,3);                           % Integration of time-sheared view

Gtu =sum(F,3);                             % Integration of time-unsheared view
```

way, CUP's forward model can be expressed by matrix multiplication, which can be computed by powerful linear algebra methods to concisely formulate solutions.

In particular, dual-view CUP's forward model, by following a matrix-vector representation, can be expressed as

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_{ts} \\ \mathbf{g}_{tu} \end{bmatrix} = \mathbf{\Phi f} = \begin{bmatrix} \mathbf{T}_{ts}\mathbf{SC} \\ \mathbf{T}_{tu} \end{bmatrix} \mathbf{f}, \qquad (3)$$

where, $\mathbf{g}_{ts} \in \mathbb{R}^{m_{ts} \times 1}$ and $\mathbf{g}_{tu} \in \mathbb{R}^{m_{tu} \times 1}$ are the vectorial version of the time-sheared view and the time-unsheared view with sizes $m_{ts} = M \cdot [N + (L-1)]$ and $m_{tu} = M \cdot N$, respectively. $\mathbf{g} \in \mathbb{R}^{m \times 1}$ is the vectorial version of the concatenated two views with a size $m = m_{ts} + m_{tu}$. $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ is the dual-view CUP's sensing matrix. $\mathbf{C} \in \mathbb{R}^{n \times n}$ is the spatial encoding matrix. $\mathbf{S} \in \mathbb{R}^{(m_{ts} \cdot L) \times n}$ is the temporal shearing matrix. $\mathbf{T}_{ts} \in \mathbb{R}^{m_{ts} \times (m_{ts} \cdot L)}$ and $\mathbf{T}_{tu} \in \mathbb{R}^{m_{tu} \times n}$ are the spatiotemporal integration matrices of the time-sheared view and the time-unsheared view, respectively. $\mathbf{\Phi}_{ts} = \mathbf{T}_{ts}\mathbf{SC} \in \mathbb{R}^{m_{ts} \times n}$ is also defined as the time-sheared sensing matrix.

The entries of $\mathbf{C}$, $\mathbf{S}$, $\mathbf{T}_{ts}$, and $\mathbf{T}_{tu}$ are given as

$$C_{i,j} = \begin{cases} r_v & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}, \qquad (4)$$

$$S_{i,j} = \begin{cases} 1 & \text{if } i = j + M \cdot L \cdot \left\lfloor \frac{j}{M \cdot N} \right\rfloor \\ 0 & \text{otherwise,} \end{cases} \qquad (5)$$

$$\mathbf{T}_{ts} = \mathbf{1}_L^T \otimes \mathbf{I}_{m_{ts} \times m_{ts}}, \qquad (6)$$

$$\mathbf{T}_{tu} = \mathbf{1}_L^T \otimes \mathbf{I}_{m_{tu} \times m_{tu}}. \qquad (7)$$

In Eq. (4), $v = \mod(j, M \cdot N)$ with $v \in \mathbb{W}$, and $r_v \in \{0,1\}$ is the value at the v'th position of $\mathbf{r} \in \mathbb{R}^{M \cdot N \times 1}$, which is the vectorial version of the encoding mask $\mathbf{R}$. In Eqs. (6) and (7), $\mathbf{1}_L \in \mathbb{R}^{L \times 1}$ is an all-one vector. $\mathbf{I}$ is the identity matrix. The matrix $\mathbf{T}_{tu}$ has a structure similar to $\mathbf{T}_{ts}$ (i.e., a horizontal concatenation of identity matrices) but with a shorter diagonal dimension.

Thus, the sensing matrix $\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_{ts} \\ \mathbf{T}_{tu} \end{bmatrix}$ can be directly defined as
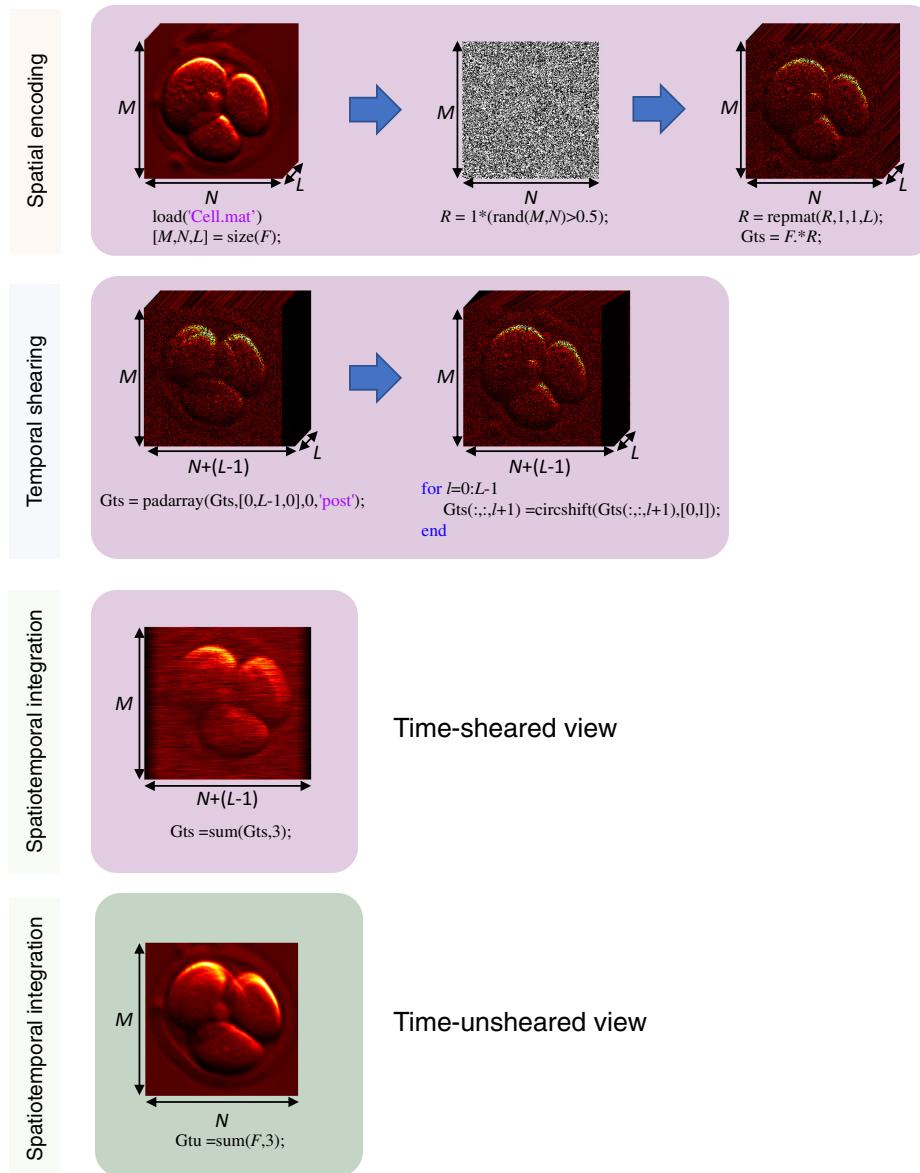
Fig. 2 Illustrations of simulating dual-view CUP's forward model with the element-wise notation using Matlab.

$$\Phi_{i,j} = \begin{cases} r_v & \text{if } i = v + M \cdot \left\lfloor \frac{j}{M \cdot N} \right\rfloor \\ 1 & \text{if } i = v + m_{\text{ts}} \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

A Matlab script for constructing the matrices $\mathbf{C}$, $\mathbf{S}$, $\mathbf{T}_{\text{ts}}$, $\mathbf{T}_{\text{tu}}$, and $\boldsymbol{\Phi}$ is presented in Algorithm 2. In this example, $M \times N \times L = 256 \times 256 \times 25$ pixels. The sensing matrix of the time-sheared view $\boldsymbol{\Phi}_{\text{ts}}$ is created by assembling a series of diagonal patterns that cyclically repeat along the horizontal direction, shifting downward by M rows following each iteration. The sensing matrix $\boldsymbol{\Phi}$, illustrated schematically in Fig. 3, is created by vertically concatenating $\boldsymbol{\Phi}_{\text{ts}}$ with $\mathbf{T}_{\text{tu}}$.

2.2 Image Reconstruction

After data acquisition, the captured 2D snapshots (i.e., $\mathbf{G}_{\text{ts}}$ and $\mathbf{G}_{\text{tu}}$) are input to an algorithm to reconstruct the dynamic scene. To date, analytical-modeling-based algorithms are dominantly used in CUP's reconstruction because they can incorporate prior knowledge about the imaging system and the underlying physics of light propagation,[69–74] leading to accurate reconstructions.

Algorithm 2 Programming matrices of dual-view CUP's operation using Matlab.

```
clear all; close all; clc

M = 256; N = 256; L = 25;

R = 1*(rand(M,N)>0.5);

R = repmat(R,1,1,L);

%%

n = M*N*L;

mts = M*(N+L-1);

mtu = M*N;

%% Encoding matrix (C)

i = 0:n-1;

j = 0:n-1;

C = sparse(i+1,j+1,R(:),n,n);

%% Shearing matrix (S)

j = 0:n-1;

i = j+(M*L)*floor(j/(M*N));

S = sparse(i+1,j+1,1,mts*L,n);

%% Integration matrices (Ts and Tu)

Tts = kron(ones(1,L),speye(mts,mts));

Ttu = kron(ones(1,L),speye(mtu,mtu));

%% CUP sensing matrix (\Phi)

Phi = [((Tts*S)*C)',Ttu']';
```

Before getting into sophisticated analytical-modeling-based reconstruction algorithms for CUP, let us analyze the structure of a basic optimization problem:

$$\tilde{\mathbf{f}} = \arg \min_{\mathbf{f}} \frac{1}{2}\|\mathbf{g} - \mathbf{\Phi}\mathbf{f}\|_2^2 + \lambda\varphi(\mathbf{f}), \tag{9}$$

where $\| \cdot \|_2^2$ is the ℓ_2-norm, $\lambda > 0$ is a regularization parameter, $\varphi(\cdot):\mathbb{R}^n \to \mathbb{R}$ is a convex and smooth function, and $\tilde{\mathbf{f}} \in \mathbb{R}^{n \times 1}$ represents the reconstruction.

Various reconstruction algorithms[75–78] are developed based on Eq. (9). A popular choice, especially in the early stage of CUP's development, is the two-step iterative shrinkage/thresholding (TwIST) algorithm.[78] The regularizer, $\varphi(\mathbf{f})$, can be set to various forms, including $\|\mathbf{\Psi}\mathbf{f}\|_1$ and $\|\mathbf{f}\|_{\text{TV}}$, where $\| \cdot \|_1$ represents the ℓ_1-norm, $\mathbf{\Psi} \in \mathbb{R}^{n \times n}$ is an arbitrary representation basis matrix, and $\| \cdot \|_{\text{TV}}$ represents the total-variation (TV) regularization.[79] The TwIST algorithm combines the shrinkage operation used in iterative soft-thresholding algorithms with a correction step that enforces fidelity to the measurements. It exploits the sparsity naturally embedded in the transient scene via the regularizer. In particular, the ℓ_1-norm requires prior knowledge about the scene to select an adequate representation basis. The TV-norm uses spatiotemporal correlation by removing low variations between neighboring pixels. These characteristics enable the TwIST algorithm to efficiently recover a transient scene from an underdetermined measurement.

Later, more advanced image reconstruction algorithms are developed based on the paradigm of alternating direction method of multipliers (ADMM),[69,80–83] which sets $\varphi(\mathbf{f}) = \|\mathbf{f} - \mathbf{f}^{(k)}\|_2^2$,
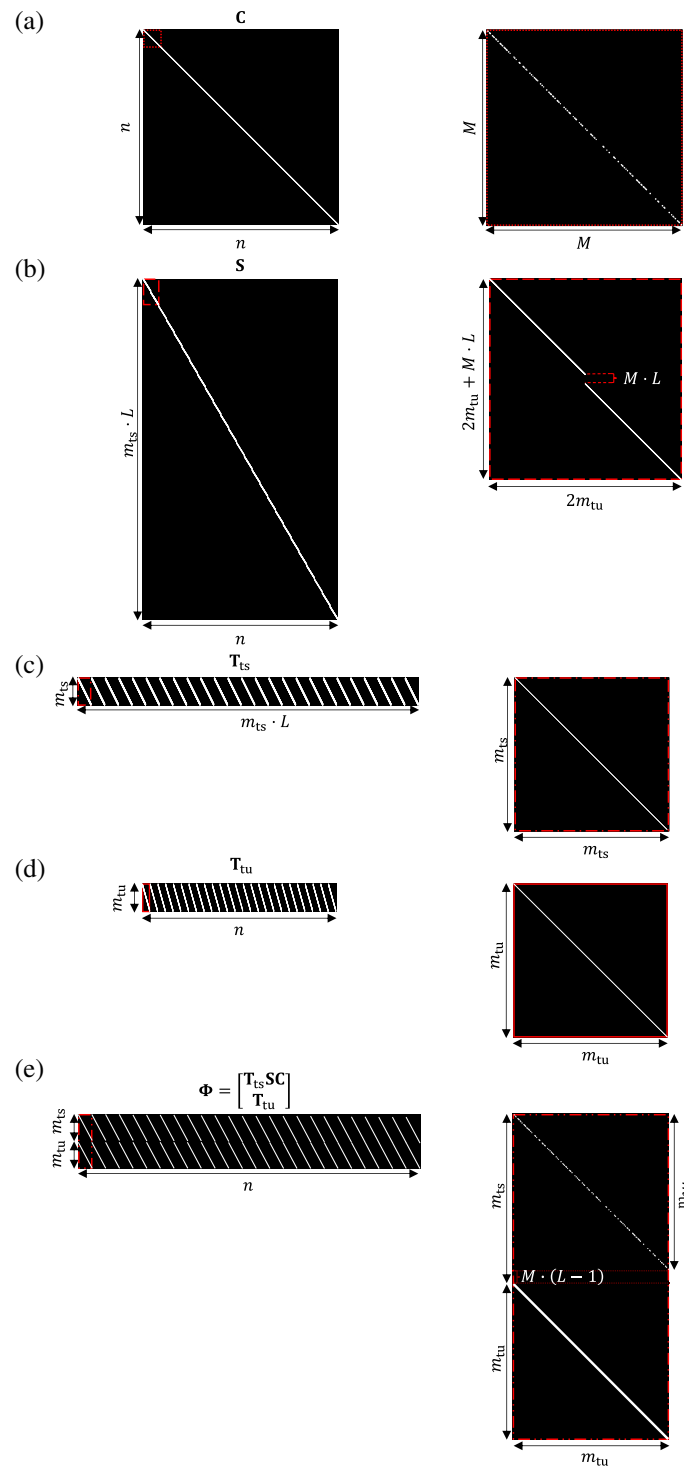
Fig. 3 Construction of the matrices of dual-view CUP's operation using Matlab. (a) Spatial encoding matrix. (b) Temporal shearing matrix. (c) Spatiotemporal integration matrix for the time-sheared view. (d) Spatiotemporal integration matrix for the time-unsheared view. (e) Sensing matrix of dual-view CUP. Insets: Zoomed-in views of local regions (indicated by the red boxes with different line types).

where $\mathbf{f}^{(k)}$ stands for the k'th reconstruction with $k = \{0, \ldots, K-1\}$ and $K \in \mathbb{N}$ as the total number of the algorithm's iterations. The ADMM has gained increasing popularity due to its flexibility to customize the optimization steps by incorporating additional constraints (e.g., noise reduction algorithms or neural-network approaches), which is hence selected for this tutorial.

2.2.1 *Analytical-modeling-based approaches*

The ADMM accomplishes distributed convex optimization using a divide-and-conquer approach, where a global problem is split into a few subproblems.[84] Leveraging dual decomposition and the augmented Lagrangian (AL) methods for constrained optimization,[77] the ADMM solves problems expressed by the following form:

$$\text{minimize } \gamma(\mathbf{f}) + \psi(\mathbf{z}) \quad \text{subject to } \mathbf{Df} + \mathbf{Bz} = \mathbf{b}, \tag{10}$$

where $\{\gamma, \psi\}$ are convex functions. In the equality constraint, $\mathbf{D}$ and $\mathbf{B}$ are arbitrary matrices that establish a linear relationship between the objective variable (i.e., $\mathbf{f}$) and an auxiliary variable (i.e., $\mathbf{z}$) that works as prior information. The variable $\mathbf{b}$ represents the limits (or bounds) for the equality constraint. For instance, in dual-view CUP, an equality constraint can be imposed by setting $\mathbf{D} = \mathbf{T}_{\text{tu}}$, $\mathbf{B} = -\mathbf{I}$, $\mathbf{z} = \mathbf{g}_{\text{tu}}$, and $\mathbf{b} = \mathbf{0}$.

Equation (10) can be solved using the method of Lagrange multipliers, which is a mathematical technique used to optimize a function subject to equality constraints. For Eq. (10), its Lagrangian function is defined as

$$\mathcal{L}(\mathbf{f}, \mathbf{z}, \boldsymbol{\nu}) = \gamma(\mathbf{f}) + \psi(\mathbf{z}) + \boldsymbol{\nu}^T(\mathbf{Df} + \mathbf{Bz} - \mathbf{b}), \tag{11}$$

where $\boldsymbol{\nu}$ is the Lagrange multiplier. As a scaling factor, $\boldsymbol{\nu}$ enables constructing, from Eq. (10), an unconstrained optimization function, in which the gradients of both the objective function and the constraint function are proportional to each other at the optimal solution.[84] Then, Eq. (11) is rewritten as[85]

$$\min_{\mathbf{f},\mathbf{z}} \max_{\boldsymbol{\nu}} \mathcal{L}(\mathbf{f}, \mathbf{z}, \boldsymbol{\nu}). \tag{12}$$

Equation (12) is maximized when $\boldsymbol{\nu} \to +\infty$ unless $\mathbf{Df} + \mathbf{Bz} - \mathbf{b} = \mathbf{0}$. By converting the maximization problem into a minimization problem [i.e., $\max_{\boldsymbol{\nu}} \mathcal{L}(\mathbf{f}, \mathbf{z}, \boldsymbol{\nu}) = \min_{\boldsymbol{\nu}} - \mathcal{L}(\mathbf{f}, \mathbf{z}, \boldsymbol{\nu})$] (Ref. 86) and using a proximal term[87] to solve the new minimization problem, Eq. (12) results in

$$\min_{\mathbf{f},\mathbf{z}} \left(\arg \min_{\boldsymbol{\nu}} - \mathcal{L}(\mathbf{f}, \mathbf{z}, \boldsymbol{\nu}) + \frac{1}{2\rho} \|\boldsymbol{\nu} - \overline{\boldsymbol{\nu}}\|_2^2 \right), \tag{13}$$

where $\rho > 0$ is the penalty parameter, and $\overline{\boldsymbol{\nu}}$ is a previous estimate of $\boldsymbol{\nu}$. Note that the "argmin" in Eq. (13) is now a convex quadratic function with the trivial solution $\boldsymbol{\nu} = \overline{\boldsymbol{\nu}} + \rho(\mathbf{Df} + \mathbf{Bz} - \mathbf{b})$. By inserting this trivial solution into Eq. (12), the AL-based dual-problem can be obtained as

$$\underset{\mathbf{f},\mathbf{z}}{\arg\min} \, \gamma(\mathbf{f}) + \psi(\mathbf{z}) + \overline{\boldsymbol{\nu}}^T(\mathbf{Df} + \mathbf{Bz} - \mathbf{b}) + \frac{\rho}{2}\|\mathbf{Df} + \mathbf{Bz} - \mathbf{b}\|_2^2. \tag{14}$$

Finally, Eq. (14) can be split into three optimization problems:

$$\mathbf{f}^{(k+1)} \triangleq \arg \min_{\mathbf{f}} \gamma(\mathbf{f}) + (\boldsymbol{\nu}^{(k)})^T(\mathbf{Df} + \mathbf{Bz}^{(k)} - \mathbf{b}) + \frac{\rho}{2}\|\mathbf{Df} + \mathbf{Bz}^{(k)} - \mathbf{b}\|_2^2, \tag{15}$$

$$\mathbf{z}^{(k+1)} \triangleq \arg \min_{\mathbf{z}} \psi(\mathbf{z}) + (\boldsymbol{\nu}^{(k)})^T(\mathbf{Df}^{(k+1)} + \mathbf{Bz} - \mathbf{b}) + \frac{\rho}{2}\|\mathbf{Df}^{(k+1)} + \mathbf{Bz} - \mathbf{b}\|_2^2, \tag{16}$$

$$\boldsymbol{\nu}^{(k+1)} \triangleq \boldsymbol{\nu}^{(k)} + \rho(\mathbf{Df}^{(k+1)} + \mathbf{Bz}^{(k+1)} - \mathbf{b}), \tag{17}$$

where $\{\boldsymbol{\nu}^{(k+1)}, \boldsymbol{\nu}^{(k)}\}$ are the equal expressions of $\{\boldsymbol{\nu}, \overline{\boldsymbol{\nu}}\}$, respectively. In this strategy, Eqs. (15)–(17) are solved in an alternating and iterative form to find a point that belongs to the intersection of the two closed convex solution sets. Here, for each step, all the parameters are fixed except the optimization variables [e.g., $\mathbf{f}$ in Eq. (15) and $\mathbf{z}$ in Eq. (16)]. Then, by repeatedly projecting the updated variables onto each set, the algorithm converges toward a point that satisfies the constraints of all the sets simultaneously.

After defining the core structure of the ADMM algorithm, the following sections discuss two popular variants of the ADMM in image processing.[84]

Scaled form ADMM. The scaled form[84] can be obtained using the equality $\boldsymbol{\nu}^T\mathbf{r} + (\rho/2)\|\mathbf{r}\|_2^2 = \frac{\rho}{2}\|\mathbf{r} + \mathbf{w}\|_2^2 - \frac{\rho}{2}\|\mathbf{w}\|_2^2$ (Ref. 86) with $\mathbf{r} = \mathbf{Df} + \mathbf{Bz} - \mathbf{b}$ and the scaled Lagrange multiplier $\mathbf{w} = \frac{1}{\rho}\boldsymbol{\nu}$. This implementation modifies Eq. (14) as

$$\arg \min_{\mathbf{f},\mathbf{z}} \gamma(\mathbf{f}) + \psi(\mathbf{z}) + \frac{\rho}{2}\|\mathbf{D}\mathbf{f} + \mathbf{B}\mathbf{z} - \mathbf{b} + \mathbf{w}\|_2^2 - \frac{\rho}{2}\|\mathbf{w}\|_2^2. \tag{18}$$

Then, setting $\mathbf{B} = -\mathbf{D} = -\mathbf{I}$ and $\mathbf{b} = \mathbf{0}$, i.e., $\mathbf{f} = \mathbf{z}$, Eq. (18) can be split into three optimization problems:

$$\mathbf{f}^{(k+1)} \triangleq \arg \min_{\mathbf{f}} \gamma(\mathbf{f}) + \frac{\rho}{2}\|\mathbf{f} - \mathbf{z}^{(k)} + \mathbf{w}^{(k)}\|_2^2, \tag{19}$$

$$\mathbf{z}^{(k+1)} \triangleq \arg \min_{\mathbf{z}} \psi(\mathbf{z}) + \frac{\rho}{2}\|\mathbf{f}^{(k+1)} - \mathbf{z} + \mathbf{w}^{(k)}\|_2^2, \tag{20}$$

$$\mathbf{w}^{(k+1)} \triangleq \mathbf{w}^{(k)} + \rho(\mathbf{f}^{(k+1)} - \mathbf{z}^{(k+1)}). \tag{21}$$

The scaled form of ADMM [i.e., Eqs. (19)–(21)] exhibits an improved convergence rate compared with the standard ADMM [i.e., Eqs. (15)–(17)]. The acceleration is achieved by introducing ρ as a scaling factor, which is particularly beneficial for large-scale optimization problems or problems with slow convergence rates. Further insights into these considerations, including heuristics for the effective selection of an appropriate scaling factor, can be found in Ref. 84.

Plug-and-play ADMM. The ADMM's modular structure is one of its main features because it enables the decomposition of a complex optimization problem [i.e., Eqs. (14) or (18)] into several simpler subproblems [i.e., Eqs. (15)–(17) or Eqs. (19)–(21)] that can be solved independently or using established solution methods. Moreover, ADMM's versatility enables modeling different sparse-based optimization problems. For example, the Tikhonov optimization problem can be modeled by setting $\psi(\mathbf{z}) = \|\mathbf{z}\|_2^2$ in Eq. (20). As another example, by setting $\mathbf{B} = \boldsymbol{\Psi}$, $\mathbf{D} = \mathbf{I}$, $\mathbf{b} = \mathbf{0}$, $\mathbf{f} = \boldsymbol{\Psi}\mathbf{z}$, and $\psi(\mathbf{z}) = \|\mathbf{z}\|_1$ in Eq. (18), Eq. (20) can be converted into the basis-pursuit denoising problem. In this regard, a popular framework is the plug-and-play (PnP)-ADMM,[69] which allows plugging in an off-the-shelf image-denoising algorithm as a solver for the subproblems (see a Matlab implementation in Algorithm 3). In the PnP-ADMM, by setting $\gamma(\mathbf{f}) = \|\boldsymbol{\Phi}\mathbf{f} - \mathbf{g}\|_2^2$, Eq. (19) has the closed-form solution

$$\mathbf{f}^{k+1} = \left[\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \frac{\rho}{2}\mathbf{I}\right]^{-1}\left[\boldsymbol{\Phi}^T\mathbf{g} + \frac{\rho}{2}(\mathbf{z} - \mathbf{w})\right]. \tag{22}$$

Then, Eq. (20) can be rewritten as a denoising problem by setting $\rho = \frac{1}{\sigma^2}$, resulting in

$$\mathbf{z}^{(k+1)} = \arg \min_{\mathbf{z}} \psi(\mathbf{z}) + \frac{1}{2\sigma^2}\|\mathbf{z} - \tilde{\mathbf{z}}^{(k)}\|_2^2, \tag{23}$$

where $\tilde{\mathbf{z}}^{(k)} = \mathbf{f}^{(k+1)} + \mathbf{w}^{(k)}$, and σ represents the denoising strength.[88] Equation (23) can be solved as

$$\mathbf{z} = \mathcal{D}_\sigma(\tilde{\mathbf{z}}^{(k)}) = \mathcal{D}_\sigma(\mathbf{f}^{(k+1)} + \mathbf{w}^{(k)}), \tag{24}$$

where $\mathcal{D}_\sigma$ is a denoiser. Note that the PnP-ADMM algorithm supports any denoiser that fulfills restrictive conditions, such as being non-expansive and having a symmetric Jacobian.[89] For example, the block-matching and 3D filtering algorithm has been extensively used to enhance the denoising capabilities of the ADMM algorithm while preserving textures and fine details.[90,91]

2.2.2 *Deep-learning approaches*

Deep-learning approaches have been increasingly featured owing to their faster reconstruction compared with their analytical-modeling-based counterparts. Recent advances have allowed embedding mathematical properties offered by the CS theory by designing custom layers that emulate the forward sensing model or exploiting spatiotemporal sparsity via image-denoising nets.[93] Given access to rich available training datasets, many novel methods based on convolutional neural networks (CNNs) have been developed for CUP's reconstruction as well as for the encoding mask design, including the end-to-end CNN with residual learning,[94] the U-Net-based DeepCUP,[95] the hybrid algorithm that combines the AL method with deep learning,[96] and the snapshot-to-video autoencoder based on a generative adversarial network.[97–103]

Algorithm 3 Simulating dual-view CUP's image reconstruction by a PnP-ADMM algorithm using Matlab[a].

```matlab
clear all

close all

clc

%% Load datacube

load('Cell.mat')

[M,N,L] = size(F);

F = F./max(F(:));

n = M*N*L;

mts = M*(N+L-1);

mtu = M*N;

m = mts + mtu;

global m

%% Mask

R = 1*(rand(M,N)>0.5);

R = repmat(R,1,1,L);

%% Sensing matrix

j = 0:n-1;

i = mod(j,M*N)+M*floor(j/(M*N));

Phi_ts = sparse(i+1,j+1,R(:),mts,n);

Phi_tu = kron(ones(1,L),speye(mtu,mtu));

Phi = [Phi_ts;Phi_tu];

%% Measurement

G = Phi*F(:);

G = G/max(G(:));

G = G/L;

%% PnP-ADMM parameters

addpath(genpath('./denoisers/RF/'));

dim = size(F);

A = @(F,trans_flag) afun2(F,trans_flag,Phi);

method = 'RF';

lambda = 0.25;

opts.rho = 0.1;

opts.gamma = 1;

opts.max_itr = 2000;

opts.print = true;

%% Main routine

F_tilde = PlugPlayADMM_general(F,G,A,lambda,method,opts,dim);
```

[a]Functions used in the tubove script can be downloaded from Ref. 92.

Here, we review a representative CNN—the deep high-dimensional adaptive net (D-HAN)[104] that offers multifaceted supervision to CUP by optimizing the encoding mask, sensing the shearing operation, and reconstructing the 3D datacubes. The main goal of the D-HAN is to leverage the merits of both the ADMM and the network-based CS methods by mapping one iteration of the ADMM steps to a deep network architecture. For these reasons, the D-HAN will be used as a benchmark to explain how to link the CUP's forward model with a CNN approach.

Originally designed to use only the time-sheared view, the D-HAN is composed of two cascaded neural networks: a deep-unfolding-based network to embody the sensing model of the time-sheared view in CUP and a U-Net architecture[105] to further improve image reconstruction (Fig. 4) by exploiting the spatiotemporal correlation of the transient scene. The deep-unfolding net and the U-Net manifest the "divide-and-conquer" approach embedded in the ADMM. Then, the time-unsheared view was incorporated to boost the reconstruction performance by using it as an initialization for the deep-unfolding network and a prior restriction in the loss function. This configuration leverages the original D-HAN's mathematical advantages and the reduction of unknowns via prior information. This design is memory efficient and thus essential for learning to reconstruct high-dimensional datacubes.

In this regard, the ADMM-based inverse problem can be formulated using the ADMM's scaled form [i.e., Eqs. (19)–(21)]. Note that in Eq. (19), the analytical inverse model of $\mathbf{f}$ refers to a quadratic problem with the closed-form that involves the inversion of a $n \times n$ size matrix [see Eq. (22)]. Toward this goal, the Sherman–Woodbury–Morrison (SWM) matrix inversion lemma[106]—a mathematical theorem allowing calculating a matrix's inverse by converting it into a full rank matrix—and the full-column rank properties are exploited to simplify the process to a smaller-scale matrix inversion and obtain the closed-form solution of the first inverse model in Eq. (22)

$$\mathbf{f} = \tilde{\rho}^{-1}[\mathbf{I} - \mathbf{\Phi}_{ts}^T[\tilde{\rho}\mathbf{I} + \mathbf{\Phi}_{ts}\mathbf{\Phi}_{ts}^T]^{-1}\mathbf{\Phi}_{ts}][\mathbf{\Phi}_{ts}^T\mathbf{g}_{ts} + \tilde{\rho}(\mathbf{z} - \mathbf{w})], \tag{25}$$

where $\mathbf{\Phi}_{ts}\mathbf{\Phi}_{ts}^T \in \mathbb{R}^{m_{ts} \times m_{ts}}$ represents a matrix product resulting in a diagonal matrix, and $\tilde{\rho} = \rho/2$.

To implement the D-HAN, the first step is to define the operators of dual-view CUP's data acquisition. First, the direct sensing operators of the time-sheared view and the time-unsheared view, denoted by $\mathcal{G}_{ts}$ and $\mathcal{G}_{tu}$, are expressed as

$$\mathcal{G}_{ts}(\mathbf{F}, \mathbf{R}) = \sum_{l=0}^{L-1} \mathcal{R}(\Gamma(l), \mathbf{F}_{:,:,l} \circ \mathbf{R}), \tag{26}$$

$$\mathcal{G}_{tu}(\mathbf{F}) = \sum_{l=0}^{L-1} \mathbf{F}_{:,:,l}. \tag{27}$$

Here, $\mathcal{G}_{ts}(\cdot) : \mathbb{R}^{M \times N \times L} \to \mathbb{R}^{M \times [N+(L-1)]}$ and $\mathcal{G}_{tu}(\cdot) : \mathbb{R}^{M \times N \times L} \to \mathbb{R}^{M \times N}$. They are shown as the magenta and orange layers in Fig. 4. $\circ$ represents the Hadamard product. The operator $\mathcal{R}(\cdot) : \mathbb{R}^{M \times N} \to \mathbb{R}^{M \times [N+(L-1)]}$ introduces a right-zero-padding (i.e., $[\mathbf{F}_{:,:,l}, \mathbf{0}]$ with $\mathbf{0} \in \mathbb{R}^{M \times (L-1)}$)
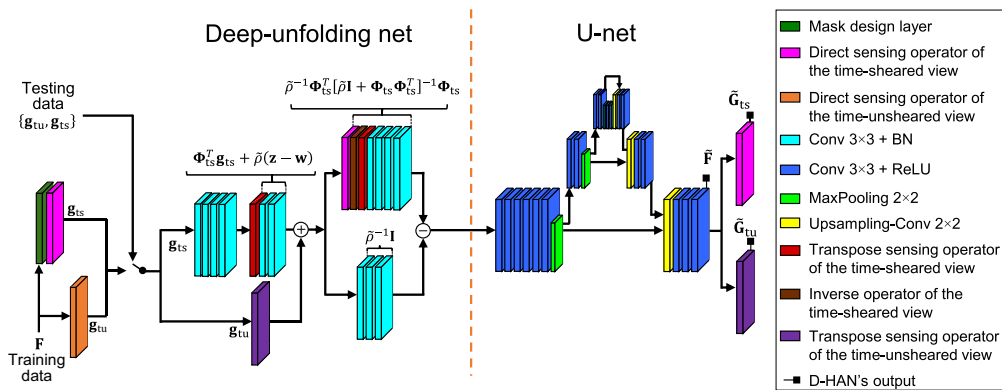

Fig. 4 Schematic of the D-HAN for dual-view CUP's image reconstruction. BN, batch normalization; ReLU, rectified linear activation unit. Adapted with permission from Ref. 104.

Algorithm 4 Programming the direct sensing operators of the time-sheared view and the time-unsheared view (i.e., $\mathcal{G}_{ts}$ and $\mathcal{G}_{tu}$) using TensorFlow.

```
## Direct sensing operator time-sheared view

class DirectSensing_ts(tf.keras.layers.Layer):

  def __init__(self, L, M, N, **kwargs):

    self.L = L

    self.M = M

    self.N = N

    super(DirectSensing_ts, self).__init__(**kwargs)

  def get_config(self):

    config = super().get_config().copy()

    config.update({

      'bands': self.L})

    return config

  def call(self, F, R, **kwargs):

    F = tf.multiply(R, F)

    F = tf.pad(F, [[0, 0], [0, 0], [0, self.L - 1], [0, 0]],name = "padsensing")

    Gts = None

    for i in range(0,self.L):

      if Gts is not None:

        Gts = Gts  + tf.roll(F[:,:,:,i], shift=i, axis=2)

      else:

        Gts = F[:,:,:,i]

    Gts = tf.expand_dims(Gts ,axis=-1)

    Gts = tf.math.divide(Gts ,self.L)

    return Gts

    ## → Output: Compressed measurement

    ## Direct sensing operator time-unsheared view

Gtu = tf.math.reduce_mean(F,axis=-1)
```

followed by a right-horizontal circular shifting of $\Gamma(l)$ pixels. A script of Python to construct the sensing operators $\mathcal{G}_{ts}$ and $\mathcal{G}_{tu}$ are presented in Algorithm 4.

Then, the transpose sensing operators of the time-sheared view and the time-unsheared view, shown as the red and purple layers respectively in Fig. 4, are defined as

$$\mathcal{F}_{ts}(\mathbf{G}_{ts}, \mathbf{R}) = \mathcal{S}(\Gamma(l), \mathbf{G}_{ts} \circ \mathcal{R}(\Gamma(l), \mathbf{R})), \tag{28}$$

$$\mathcal{F}_{tu}(\mathbf{G}_{tu}) = (\mathbf{G}_{tu})_{:,:,l}. \tag{29}$$

Here, $\mathcal{F}_{ts}(\cdot)\colon \mathbb{R}^{M \times [N+(L-1)]} \to \mathbb{R}^{M \times N \times L}$ and $\mathcal{F}_{tu}(\cdot)\colon \mathbb{R}^{M \times N} \to \mathbb{R}^{M \times N \times L}$. They return a data-cube from a 2D compressed measurement. $\mathcal{S}(\cdot)$ is an operator that performs a left-horizontal circular shifting of $\Gamma(l)$ pixels, followed by the removal of the last $(L-1)$ columns in the

Algorithm 5 Programming the transpose sensing operators of the time-sheared view and the time-unsheared view (i.e., $\mathcal{F}_{ts}$ and $\mathcal{F}_{tu}$) using TensorFlow.

```
## Transpose sensing operator of the time-sheared view

class Transposesensing_ts(tf.keras.layers.Layer):

  def __init__(self, L, M, N, **kwargs):

    self.L = L

    self.M = M

    self.N = N

    super(Transposesensing_ts, self).__init__(**kwargs)

  def get_config(self):

    config = super().get_config().copy()

    config.update({

      'bands': self.L})

    return config

  def call(self, Gts, R, **kwargs):

    F = None

    R = R[0,:,:,0]

    Gts = Gts[:,:,:,0]

    for i in range(0,self.L):

      if F is not None:

        Ab = tf.roll(Gts, shift=-i, axis=2)

        Ax = tf.expand_dims(tf.multiply(R, Ab[:,:,0:self.N]), -1)

        F = tf.concat([[F, Ax], axis=-1)

      else:

        Ab = tf.roll(Gts, shift=0, axis=2)

        F = tf.expand_dims(tf.multiply(R,Ab[:,:,0:self.N]), -1)

    F = self.L*F

    return F

## Transpose sensing operator of the time-unsheared view

Gtu = tf.expand_dims(Gtu,axis=-1)

F_tu = tf.broadcast_to(Gtu, [Gtu.shape[0], M, N, L])
```

resulting shifted matrix to preserve the spatial dimension of the datacube. Algorithm 5 presents a Python script to construct $\mathcal{F}_{ts}$ and $\mathcal{F}_{tu}$.

Finally, the inverse operator of the time-sheared view, shown as the brown layer in Fig. 4, is defined as

$$\mathcal{I}_{ts}(\mathbf{G}_{ts}, \mathbf{R}) = \mathbf{G}_{ts} \circ \left(\sum_{l=0}^{L-1} \mathcal{R}(\Gamma(l), \mathbf{R}^{\circ 2}) + \tilde{\rho}\mathbf{I} \right)^{\circ -1}, \qquad (30)$$

where $\mathcal{I}_{ts}(\cdot):\mathbb{R}^{M\times[N+(L-1)]} \rightarrow \mathbb{R}^{M\times[N+(L-1)]}$. $(\cdot)^{\circ 2}$ and $(\cdot)^{\circ -1}$ represent the Hadamard quadratic power and the Hadamard inverse operation, respectively. An example script of Python to construct the inverse operator of the time-sheared view $\mathcal{I}_{ts}$ is summarized in Algorithm 6.

Algorithm 6 Programming the inverse operator of the time-sheared view (i.e., $\mathcal{I}_{ts}$) using TensorFlow.

```python
class InverseOperator_ts(tf.keras.layers.Layer):

  def __init__(self, L, M, N, **kwargs):

    self.L = L

    self.M = M

    self.N = N

    super(InverseOperator_ts, self).__init__(**kwargs)

  def get_config(self):

    config = super().get_config().copy()

    config.update({

      'bands': self.L})

    return config

  def build(self, input_shape):

    Lambda = tf.constant_initializer(1)

    Tau = tf.constant_initializer(1)

    Psi = np.zeros([self.M,self.N+self.L-1])

    Psi = tf.constant_initializer(Psi)

    self.Lambda = self.add_weight(name="Lbd," initializer=Lambda,
shape=(1),trainable=True)

    self.Tau = self.add_weight(name="Tau," initializer=Tau, shape=(1),
trainable=True,constraint=tf.keras.constraints. MaxNorm(max_value=1, axis=0))

    self.Psi = self.add_weight(name="Psi," initializer=Psi, shape=(self.M,self.N
+self.L-1),trainable=True)

    super(InverseOperator_ts, self).build(input_shape)

  def call(self, Gts, R, **kwargs):

    Gts = Gts[:,:,:,0]

    R1 = tf.broadcast_to(R,[1,self.M,self.N,self.L])

    Gp = DirectSensing_ts(L=self.L, M=self.M, N=self.N,
name='DirectPr_InitInv')(R, R1)

    Gp = Gp[:,:,:,0]

    Gp = Gp/(self.Lambda) + tf.ones(Gp.shape)

    Inv = tf.math.reciprocal(Gp, name=None)

    Gts = tf.multiply((self.Tau**2)*Inv+(1-self.Tau**2)*self.Psi,Gts)

    Gts = tf.expand_dims(Gts,axis=-1)

    F = TransposeSensing_ts(L=self.L, M=self.M,N=self.N,
name='TransPr_InitInv')(Gts,R)

    #

    F = F/(self.Lambda**2)

    return F
```

Following the definition of these five operators, the next step is to model the SWM matrix approach. Toward this goal, Eq. (25) is split into two main equations $\mathbf{\Phi}_{ts}^T \mathbf{g}_{ts} + \tilde{\rho}(\mathbf{z} - \mathbf{w})$ and $\tilde{\rho}^{-1}\mathbf{I} - \tilde{\rho}^{-1}\mathbf{\Phi}_{ts}^T[\tilde{\rho}\mathbf{I} + \mathbf{\Phi}_{ts}\mathbf{\Phi}_{ts}^T]^{-1}\mathbf{\Phi}_{ts}$. In the D-HAN, the first equation is reflected as $\mathcal{F}_{ts}$ coupled to two 2D convolutional layers, each of which with a batch normalization operation (referred to hereafter as a 2D convolutional + batch normalization (BN) layer and shown in cyan in Fig. 4). The output from the 2D convolutional + BN layer is added with an estimate from the time-unsheared view generated by $\mathcal{G}_{tu}$ and $\mathcal{F}_{tu}$. Subsequently, the second equation is represented by two parallel arms. The upper arm, corresponding to $\tilde{\rho}^{-1}\mathbf{\Phi}_{ts}^T[\tilde{\rho}\mathbf{I} + \mathbf{\Phi}_{ts}\mathbf{\Phi}_{ts}^T]^{-1}\mathbf{\Phi}_{ts}$, is composed of $\mathcal{G}_{tu}$ as a first layer followed by $\mathcal{I}_{ts}$ and $\mathcal{F}_{ts}$ along with four 2D convolutional + BN layers. The bottom arm, which corresponds to $\tilde{\rho}^{-1}\mathbf{I}$, has three 2D convolutional + BN layers. The outputs of both arms are subtracted and given as the input to the U-Net in the D-HAN that reflects Eq. (24). In the U-Net, the datacube passes through an encoding pathway comprised of max-pooling layers that simultaneously reduce the spatial dimension and increase the channels. This down step returns a smaller-size datacube with the more meaningful high-level details of the image (e.g., edges, textures, or shapes) linked to the scene's sparsity. Then, in the decoding step, comprised of upsampling layers, the U-Net reconstructs the full-size datacube (denoted by $\tilde{\mathbf{F}}$) using these learned high-level details.

The loss function $\mathcal{L}(\cdot)$, used to learn the D-HAN's weights, is established as

$$\mathcal{L}(\mathbf{F}) = l_1(\mathbf{F}, \tilde{\mathbf{F}}) + l_1(\mathbf{G}_{ts}, \tilde{\mathbf{G}}_{ts}) + l_1(\mathbf{G}_{tu}, \tilde{\mathbf{G}}_{tu}) + l_{SSIM}(\mathbf{F}, \tilde{\mathbf{F}}), \tag{31}$$

where $\tilde{\mathbf{F}}$ is the D-HAN's output, $\tilde{\mathbf{G}}_{ts}$ and $\tilde{\mathbf{G}}_{tu}$ are estimations of the compressed measurement from $\tilde{\mathbf{F}}$ using Eqs. (26) and (27), respectively. $l_1(\cdot)$ is the l_1-norm operator, and $l_{SSIM}(\cdot)$ represents the structural similarity (SSIM) index.[107]

2.2.3 Simulation

CUP's image reconstruction of the "cell-division" scene is simulated using both the analytical-modeling-based algorithm (in Matlab) and deep-learning algorithm (in Python). The dimensions of the datacube were set as $M \times N \times L = 256 \times 256 \times 25$ pixels, and the binary mask holds the structure proposed in Ref. 104. Four popular databases—"SumMe,"[108] "Need for Speed,"[109] "Sports Videos in the Wild,"[110] and "Mouse Embryo Tracking"[67]—were used to train the D-HAN. The PnP-ADMM algorithm and a pretrained version of the D-HAN can be downloaded from Ref. 92 (Matlab 2022b) and Ref. 111 (Python, TensorFlow). In addition, a more beginner-friendly Python version is available in Ref. 112, which was trained on the Google Colaboratory (CoLab) application—a free Jupyter Notebook interactive development environment for Python hosted in Google's cloud.

Six exemplary frames of the scene (as the ground truth) and their corresponding frames reconstructed by single-view and dual-view CUP using the PnP-ADMM and the D-HAN are shown in Fig. 5(a). The movie is shown in Video 1. As shown in Figs. 5(b) and 5(c), results show that implementing the dual-view approach exceeds the reconstruction performance of a single-view CUP in terms of the average peak signal-to-noise ratio (PSNR) defined as $\overline{\text{PSNR}} = \frac{1}{L}\sum_{l=0}^{L-1}\left[10\log_{10}\left(\frac{[\max(\mathbf{F}_{:,:,l})]^2}{m_{tu}^{-1}\|\text{vec}(\mathbf{F}_{:,:,l}) - \text{vec}(\hat{\mathbf{F}}_{:,:,l})\|_2^2}\right)\right]$ and the average SSIM index[113] defined as $\overline{\text{SSIM}} = \frac{1}{L}\sum_{l=0}^{L-1}[[\text{Lum}(\mathbf{F}_{:,:,l}, \hat{\mathbf{F}}_{:,:,l})]^\alpha[\text{Cont}(\mathbf{F}_{:,:,l}, \hat{\mathbf{F}}_{:,:,l})]^\beta[\text{Struc}(\mathbf{F}_{:,:,l}, \hat{\mathbf{F}}_{:,:,l})]^\gamma]$. Here, $\text{vec}(\cdot)$ is a vectorization operator, and $\hat{\mathbf{F}}$ is the reconstructed result. The operators $\text{Lum}(\cdot) = \frac{2\mu_x \cdot \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$, $\text{Cont}(\cdot) = \frac{2\sigma_x \cdot \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$, and $\text{Struc}(\cdot) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$ measure the similarities in luminance, contrast, and structure, respectively, where $\mu_x, \mu_y, \sigma_x, \sigma_y$, and σ_{xy} are the local means, standard deviation, and cross-covariance for the images. $\{\alpha, \beta, \gamma\} > 0$ are parameters used to adjust the relative importance of the three components. C_1, C_2, and C_3 are constants to stabilize the division with weak denominator. For the results shown in Fig. 5, SSIM's parameters were set as $\alpha = \beta = \gamma = 1$, $C_1 = 0.01^2$, $C_2 = 0.03^2$, and $C_3 = C_2/2$. The D-HAN obtains a better average PSNR and a comparable average SSIM to the PnP-ADMM approach in both single-view and dual-view CUP.
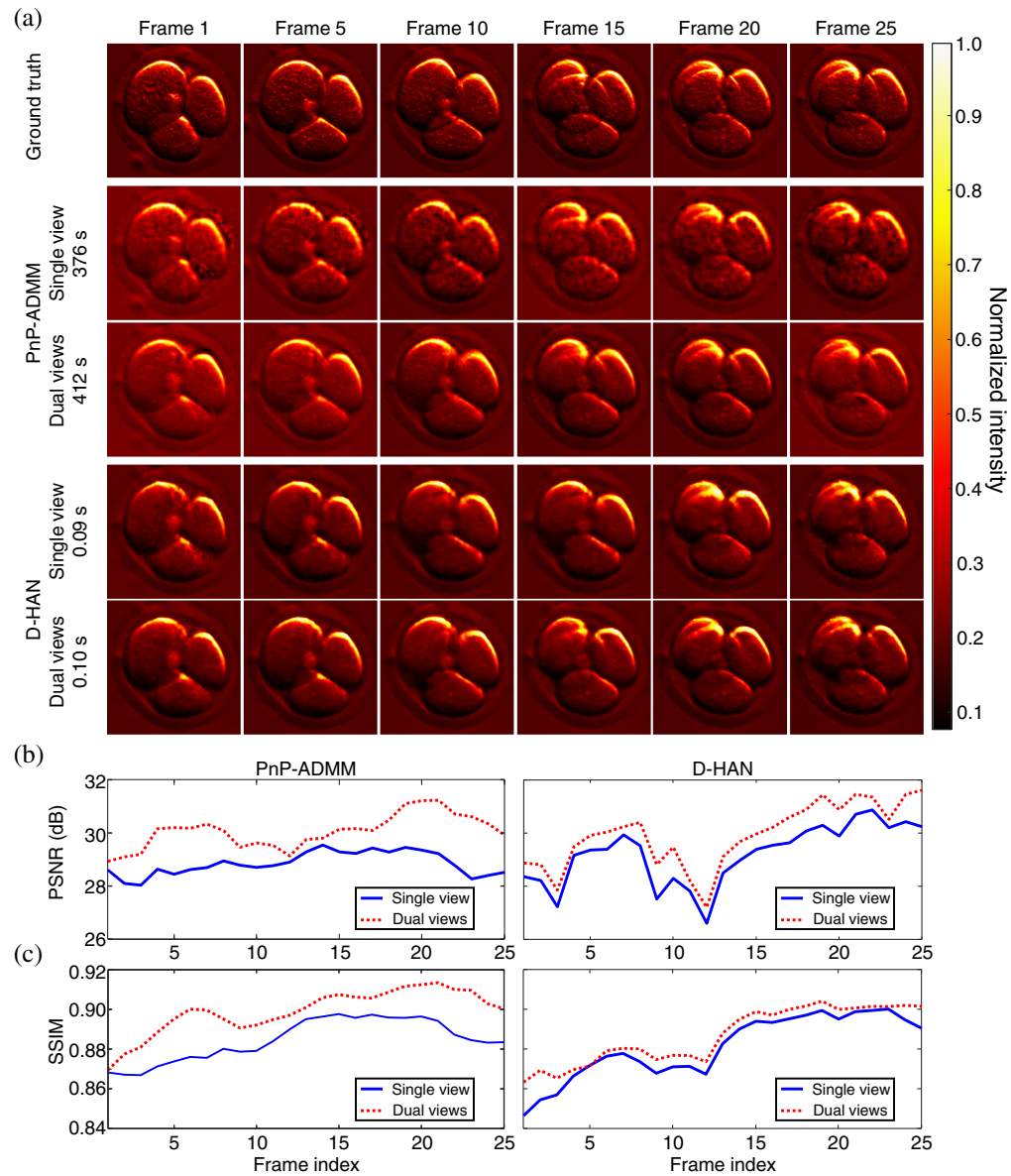
Fig. 5 Simulation of CUP's image reconstruction of the "cell-division" dynamic scene using the PnP-ADMM algorithm and the D-HAN. (a) Six selected frames of the ground truth and the single-view and dual-view reconstructions using the PnP-ADMM and D-HAN algorithms. (b) PSNR of each reconstructed frame. (c) As (b), but showing the SSIM index (Video 1, MP4, 408 KB [URL: https://doi.org/10.1117/1.JBO.29.S1.S11524.s1]).

CUP's performance decreases with higher noise and stronger compression. Table 1 illustrates this general trend from an ablation analysis of the "cell-division" dynamic scene using the PnP-ADMM algorithm. Higher noise levels reduce spatial resolution. Higher compression ratios result in stronger blurring in the temporal shearing direction, which further decreases the spatial resolution in that direction.[59,91,114] Both factors hamper the reconstruction algorithm's ability to accurately place the correct amount of intensity from the compressed snapshot to the appropriate spatiotemporal position in the reconstructed datacube.

3 System

The construction of a CUP system involves a careful selection of three crucial components. First, a spatial encoder modulates the dynamic event. Second, a temporal shearing unit deflects the spatially encoded frames to different spatial positions according to their time of arrival.

Table 1 Average PSNRs in reconstructed datacubes with different compression ratios and signal-to-noise ratios (SNRs).

Compression ratio	SNR (dB)				
	15	20	25	30	Infinity
10.1×	26.3 ± 0.8	27.9 ± 0.5	29.5 ± 0.7	30.0 ± 0.8	30.4 ± 0.9
11.9×	26.2 ± 0.6	27.7 ± 0.5	29.2 ± 0.6	29.9 ± 0.7	30.2 ± 0.8
16.4×	26.1 ± 0.5	27.5 ± 0.5	28.7 ± 0.5	29.7 ± 0.7	29.9 ± 0.7
22.9×	25.8 ± 0.8	27.4 ± 0.4	28.2 ± 0.6	28.8 ± 0.7	28.9 ± 0.8
45.6×	25.6 ± 0.9	27.1 ± 0.8	28.1 ± 0.9	28.6 ± 1.0	28.8 ± 1.1

Finally, a 2D sensor integrates the spatially encoded and temporally sheared datacube into the time-sheared view. For dual-view CUP, another 2D sensor integrates the dynamic scene into the time-unsheared view. To date, many approaches have been implemented to devise each component. A comprehensive survey of these implementations, their advantages, and limitations will be presented in this section, followed by a discussion on important steps to calibrate a CUP system.

3.1 Spatial Encoder

The selection of a suitable spatial encoder in a CUP system includes encoding pattern design and the encoder's implementation. Because CUP relies on CS principles, its sensing matrix can be designed based on the restricted isometry property (RIP) to ensure its incoherence to the representation matrix of the scene. Notably, the sensing matrix based on a random pattern has been verified to meet the RIP criterion for a wide range of representation bases.[115] Therefore, pseudorandom masks [Fig. 6(a)] are dominantly implemented as spatial encoders in reported CUP systems.

The RIP also provides a valuable metric for evaluating the encoder's quality. The general strategy is to reduce the coherence between the sensing matrix and the representation matrices to ensure that the projection of high-dimensional data onto a lower-dimensional space preserves the essential data features.[116,117] It guarantees that the compressed measurements retain sufficient information to accurately reconstruct the original signal. To date, several works have improved the mask via deep learning.[104,118,119] As an example, an encoding mask designed via the D-HAN is shown in Fig. 6(b), where the shearing operation in CUP's forward model and the training data produce horizontal stripe-like structures.[104]

As summarized in Table 2, four approaches have been used to implement CUP's spatial encoders: digital micromirror devices (DMDs),[17,66,102,120,121] liquid-crystal spatial light modulators (LC-SLMs),[134] high-definition printing,[97,135] and photolithography.[91] Among them, the DMD, as a reflective binary-amplitude spatial light modulator,[136] can provide reconfigurable, stable, and broadband encoding [Fig. 7(a)]. However, due to the micromirror's tilt angle, the
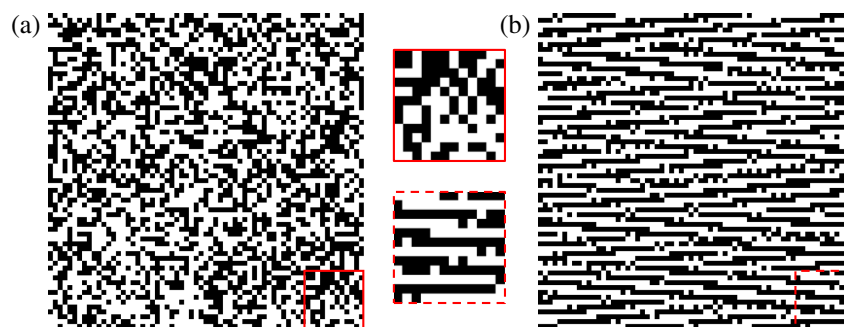

Fig. 6 Representative encoding patterns for CUP. (a) Pseudorandom pattern. (b) Deep-learning-optimized pattern. Insets: Zoomed-in views of local regions.

Table 2 Representative approaches for CUP's spatial encoders.

| Approach	Advantage	Limitation	References
DMD	— Programmable encoding — Broad operating spectrum	— Restricted FOV due to the Littrow configuration — Energy loss due to the limited fill factor and diffraction — Nonoptimal spectral response to the micromirror's coating	17, 55, 57, 58, 60–62, 64–66, 96, 99, 102, 104, 114, 116, 120–133
LC-SLM	— Programmable encoding — Phase and amplitude modulation in grayscale — Reflective and transmissive encoding ability	— Wavelength and polarization sensitive modulation — Relatively low fill factor for the transmissive type — Flicker noise	134
High-definition printing	— Transmissive encoding — Low cost — Broad operating spectrum	— Unreconfigurable encoding	97, 135
Photolithography	— Transmissive encoding — High resolution — Broad operating spectrum	— Unreconfigurable encoding — High cost	91
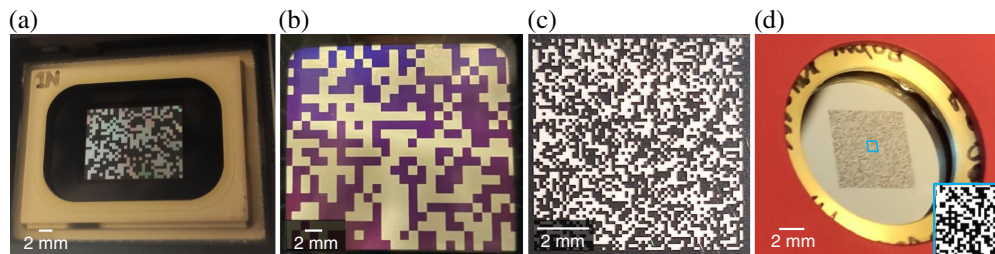
Fig. 7 Pseudorandom binary masks displayed on representative spatial encoders. (a) DMD. (b) LC-SLM. (c) Plastic mask fabricated by high-definition printing. (d) Chromium mask made by photolithography. Inset: Zoom-in view of a local region.

DMD is often required to be placed in the Littrow configuration in CUP systems[17,37,59] to retro-reflect the incident light. Since the DMD is not parallel to the object plane, this design limits the field of view (FOV). Moreover, the DMD's structure limits light efficiency in three main aspects. First, since the DMD has a ~94% fill factor, a part of incident light is lost in the gaps between neighboring micromirrors. Second, as a 2D diffraction grating,[136] the DMD has an overall diffraction efficiency of ~86%,[137] which indicates energy loss in high-diffraction orders. Finally, the aluminum coating of the micromirrors has a reflectivity of 89% in the visible spectrum with a dip at around 800 nm corresponding to the absorption of inter-band transitions in aluminum. Consequently, the constructed CUP system may not have an optimal spectral response for the dynamic scenes under investigation.

Another choice of reconfigurable spatial encoding is LC-SLMs. They have been widely implemented in coded optical imaging.[18,138–141] LC-SLMs can simultaneously modulate amplitude and phase in grayscale.[134] In the context of CUP, they can provide both reflective and transmissive spatial encoding [Fig. 7(b)]. Nonetheless, LC-SLMs could also bring in some limitations in spatial encoding. Its modulation is sensitive to both the wavelength and polarization.

Moreover, a relatively low fill factor of transmissive LC-SLMs (e.g., 58%)[142] and the flicker noise could limit pattern quality and encoding stability.[143]

Besides using the programmable devices, an encoded mask can be directly fabricated on a substrate. As a representative approach, high-definition printing can manufacture encoding masks at up to 50,800 dots per in. resolutions, with up to ∼30 in. × 30 in. in size, at $16.7 per in.^{2} (Ref. 144) [Fig. 7(c)]. In one printing task, users can pack multiple masks with different encoding pixel sizes down to 7 μm and different pattern types as well as calibration patterns, such as single pinholes, pinhole arrays, and slits. As another approach, photolithography can produce spatial encoders with nanometer-level encoding pixel sizes over inches [Fig. 7(d)]. As an example, a 3-in. × 3-in. mask with 125-nm resolution can be fabricated at ∼$6,000. As a well-established fabrication technique, photolithography can be used with various materials to target different spectral bands.[145] These fabricated coded masks can be directly inserted in CUP systems, which conserves space for a more compact system design. Although capable of providing broadband and transmissive encoding, these two approaches can only prepare fixed spatial encoders. In addition, the almost unavoidable defect pixels in the fabricated encoder request careful calibration to build an accurate sensing matrix.

3.2 Temporal Shearing Unit

Depending on the necessity of external power, temporal shearing units can be classified into passive units and active units (Table 3). The former deflects the temporal information transferred to certain photon tags (e.g., wavelengths) by exploiting the properties in these tags (e.g., color dispersion). Being jitter-free, these compact units bring in stable operation without increasing the control complexity.[59] The active units are driven by time-varying electric signals to trigger deflection. Usually integrated into the detection side of the imaging systems, they enable receive-only detection, which is specifically suited for capturing self-luminescent and color-selective events.[16,18,66]

Figure 8 shows two examples of passive temporal shearing units. Both need to team up with a chirped ultrashort probe pulse, which maps the temporal information of the event to its spectral band. As shown in Fig. 8(a), the modulated chirped pulse is spatially dispersed by a grating.[59] In recent years, the development of metamaterials has made metalenses a potential passive temporal shearing unit. They consist of an array of waveguide structures with a subwavelength size, with resonant metamaterial elements etched into the surface [Fig. 8(b)].[152,153] Metalenses can strongly disperse light while manipulating its phase, amplitude, and polarization.[154] This property has been exploited in hyperspectral imaging.[155] Grafting this sensing paradigm in CUP, a metalens integrates imaging and temporal shearing, which greatly reduces the system's size and complexity.[146] Besides the aforementioned two units, other dispersive optical elements such as kinoforms,[156] zone plates,[157] and diffractive optical elements (DOEs)[158] could also be used for passive temporal shearing of chirped pulses.

Active temporal shearing units have also been featured in many CUP systems. As an example, the image-converter streak tube is shown in Fig. 9(a). Such a device works by directing the dynamic scene onto a photocathode, where the incident photons are converted to photoelectrons. After being accelerated by a pulling voltage added on a metal mesh, these photoelectrons are temporally sheared by a varying electric field produced by applying a voltage to a pair of sweep electrodes. Then, the photoelectric signal is amplified by a microchannel plate. Finally, the photoelectrons bombard a phosphor screen and are converted back to photons.[59,91] The configuration of the image-converter streak tube takes advantage of the movement of electrons under high-voltage electric fields, enabling ultrafast shearing for the CUP system to provide up to femtosecond-level temporal resolution.[58,59] However, this operation is inevitably affected by the electronic jitter. Moreover, due to the space-charge effect in electronic imaging,[160] a trade-off needs to be made between the incident light intensity and the signal gain, which limits the imaging quality of the streak tube-based CUP systems.[37,151] The efficiency of image-converter streak tubes is also inherently limited by the photon–electron–photon conversion. The quantum yield of the photocathode is moderate for the visible light and decreases dramatically for near-infrared light.[161,162] The phosphor screen also has a relatively low conversion efficiency, especially for the fast-responding types.[163] The limited overall efficiency makes the image-converter streak tubes less suitable for imaging faint transient events.

Table 3 Representative methods for CUP's temporal shearing units.

Category	Approach	Advantage	Limitation	References
Passive	Grating	— Compact	— Requirement of chirped pulse illumination	29, 59, 121
		— Low cost		
		— Jitter-free	— Fixed shearing rate	
		— Ultrafast shearing		
	Metalens	— Compact, lightweight, and less complex optomechanically	— Requirement of chirped pulse illumination	146
		— Joint temporal shearing and imaging	— Fixed shearing rate	
		— Jitter-free	— Limited aperture size	
		— Ultrafast shearing	— High cost	
Active	Image-converter streak tube	— Receive-only detection	— High cost	17, 55, 57, 58, 60–62, 64–66, 91, 96, 99, 102, 104, 114, 116, 120, 121, 123–131, 147
		— Tunable shearing speeds	— Space-charge effect	
		— Ultrafast shearing	— Electronic jitter	
			— Low overall efficiency	
			— Spectra limited by the photocathode	
	Rotating mirror	— Receive-only detection	— Relative slow shearing speed	37, 97, 148, 149
		— Tunable shearing speeds		
		— All-optical operation		
		— Broad operating spectrum		
		— Low cost		
	TDI technique	— Receive-only detection	— Fixed shearing speed	134, 150
		— Joint temporal shearing and spatiotemporal integration	— Relative slow shearing speed	
	Electro-optical deflector	— Receive-only detection	— Small numerical aperture	122
		— All-optical operation	— High operating voltage	
		— Ultrafast shearing	— Limited deflection angle	
	Molecular deflector	— Receive-only detection	— Requirement of an ultrafast, high-intensity pump laser pulse	151
		— All-optical operation		
		— Small size		
		— Ultrafast shearing		

Rotating mirrors are another popular choice of active temporal shearing units for CUP. The mirror rotation continuously alters the angle of incidence, hence shearing the reflected light [Fig. 9(b)]. Rotating mirrors are preferred to be placed at the Fourier plane of a $4f$-system so that after the second lens, the chief rays of all temporal frames can propagate and enter the
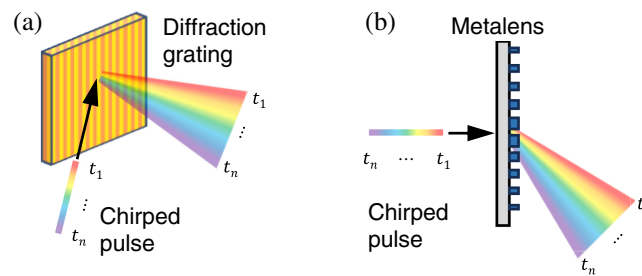
Fig. 8 Representative passive temporal shearing units for CUP. The temporal information is mapped to the spectrum and deflected to different spatial positions by (a) a diffraction grating and (b) a metalens. t_1 to t_n, temporal information.
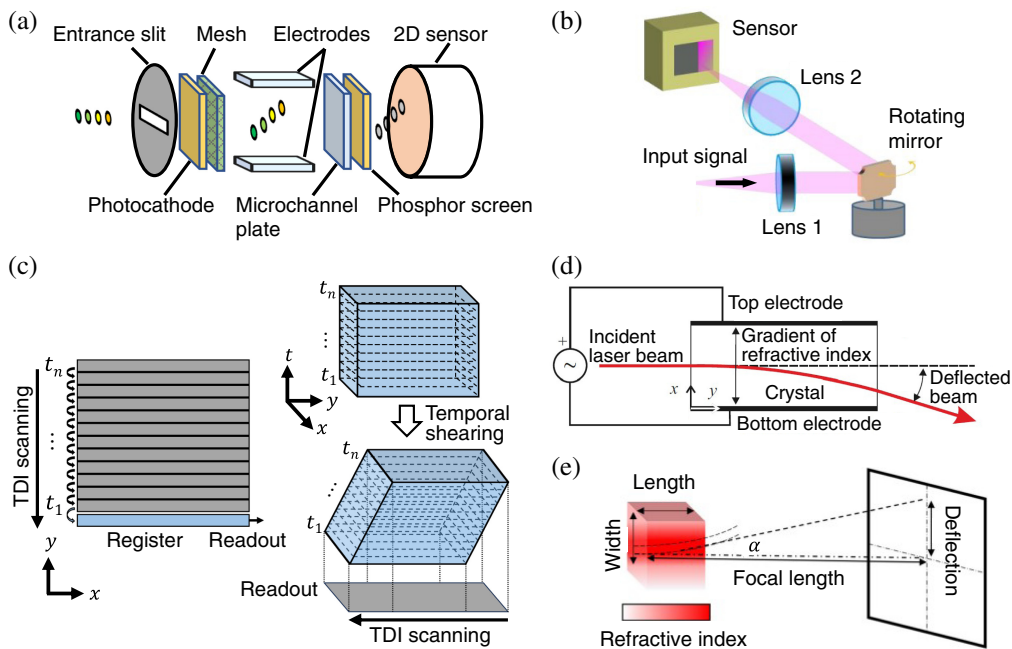

Fig. 9 Representative active temporal shearing units for CUP. (a) Image-converter streak tube. (b) Rotating mirror. (c) TDI mode of a CCD camera. (d) Electro-optical deflector. (e) Ultrashort-pulse-induced CO_2 molecule deflector. α, Deflection angle. (b) Reprinted with permission from Ref. 135. (d) Adapted with permission from Ref. 159. (e) Adapted with permission from Ref. 151.

sensor perpendicularly, which avoids aberrations introduced by the field curvature.[164] Producing tunable temporal resolutions typically from hundreds of nanoseconds to microseconds, they are much slower than the image-converter streak tube. However, the all-optical operation avoids the space-charge effect, which enables optics-limited spatial resolution and high dynamic ranges.[37] Moreover, by circumventing the photon-to-photoelectron conversion in a photocathode, rotating-mirror-based CUP systems can employ sensors in matching responsive bands to sense photons with relatively low energy (e.g., in the infrared range). Leveraging high reflectivity coatings (e.g., >95% at 0.4 to 20.0 μm),[165] these CUP systems are attractive candidates for high-sensitivity transient imaging at broad spectral bands.

Besides these two popular approaches, other specialized optical and/or electronic devices have been implemented as CUP's temporal shearing units. As an example, Fig. 9(c) shows the operation principle of the time-delay-integration (TDI) mode of a CCD camera. Initially developed to visualize moving objects under extremely low light levels, the TDI configuration employs a long exposure during which the generated photoelectrons shift down row by row before eventually reading out.[166] In this way, the read-out data are the integration of information from different rows at different time points. Such a mechanism enables TDI cameras to combine the operations of temporal shearing and spatiotemporal integration, which considerably reduces the system's

complexity.[134,150] Recently, electron-transfer-based temporal shearing has also been implemented in a streak-camera sensor.[167,168] Hundreds of sampling and storage cells are placed underneath a line of photodiodes. During the sensor's exposure, the 1D signal is sampled and sequentially stored at a temporal resolution of 500 ps. Although its 1D FOV excludes its implementation with CUP, this highly integrated device marks its potential to be further developed for future CUP systems.

Electro-optic crystals can also be used as the temporal shearing unit of CUP systems. As shown in Fig. 9(d), a time-varying electric field is applied to modulate the gradient of the refractive index of an electro-optic crystal. In this way, this electro-optic deflector (EOD) can direct the incident light to different propagation directions according to its time of arrival.[159,169] The EOD is currently the only all-optical shearing unit capable of achieving 50×10^9 frames per second (fps) in a CUP system.[122] However, the shortcomings of small numerical aperture, high operating voltage, and limited deflection angle still hinder EODs for further applications in CUP.

Finally, transient materials' behaviors have been proposed as CUP's temporal deflectors. Figure 9(e) depicts how the transient alignment of CO_2 molecules excited by an ultrashort laser pulse can induce a time-varying refractive-index gradient, resulting in different deflection angles to temporally shear the dynamic scene.[151] Although having not been experimentally demonstrated, this mechanism could open a new avenue of transient-event-assisted ultrafast imaging. The fast responses of properly selected materials could push CUP's imaging speed to the quadrillion fps level.[170]

3.3 2D Sensor

After being spatially encoded and temporally sheared, the dynamic scene is spatiotemporally integrated over each pixel by a 2D sensor. Most of the current commercial cameras (e.g., CCD, CMOS, scientific CMOS, and electron-multiplying CCD cameras) have been implemented to construct a CUP system. Nonetheless, used as the last component in a CUP system, the 2D sensor needs to be carefully selected to accommodate the characteristics of the dynamic scenes, the spatial encoders, and the temporal shearing units. In terms of spectral responsiveness, the 2D sensors are desired to have the highest sensitivity at the corresponding spectra of the dynamic scene. However, it might be restricted by the device. For example, for the image-converter streak tube, the quantum yield of the deployed camera should be peaked at the wavelength of the phosphor screen (e.g., 540 nm of a P43 phosphor screen). The pixel size of these sensors is required to sufficiently sample each encoding pixel for the given system's magnification.

The shutter type is another important factor in the sensor selection. Overall, the global shutter is much preferred for CUP operation compared with the rolling shutter. Figure 10(a) shows a simulated dynamic scene of a rotating spinner with constant intensity. For a rolling-shutter sensor, the exposure of each row starts sequentially from the top to the bottom for the same period and ends at a different time point. The induced rolling-shutter effect distorts the image of fast-moving objects.[171] CUP can overcome this distortion by putting the information back in the correct spatiotemporal position. However, due to the different starting times of exposure, only a part of FOV can be reconstructed for images at the beginning and the end of the movie, as shown in Fig. 10(b). This issue can be bypassed by limiting the occurrence of the dynamic event when all rows are under exposure. In contrast, the global shutter, which can be implemented in both CCD and CMOS sensors, allows capturing the dynamic scene over the full FOV [Fig. 10(c)] and thus avoid the time-windowing effect of the rolling shutter.

3.4 Calibrations

In this section, we outline a few important calibration steps in CUP's operation. They are necessary for both physical data acquisition and computational reconstruction of the dynamic scene.

3.4.1 Co-registration of multiple views

Due to the difference among individual imaging arms, the acquired snapshots may have different aberrations. It is thus indispensable to accurately co-register all the views for accurate image reconstruction. Toward this goal, a static image of the time-sheared view is acquired by turning off the shearing unit (Fig. 11). In Matlab, the co-registration can be carried out using "control
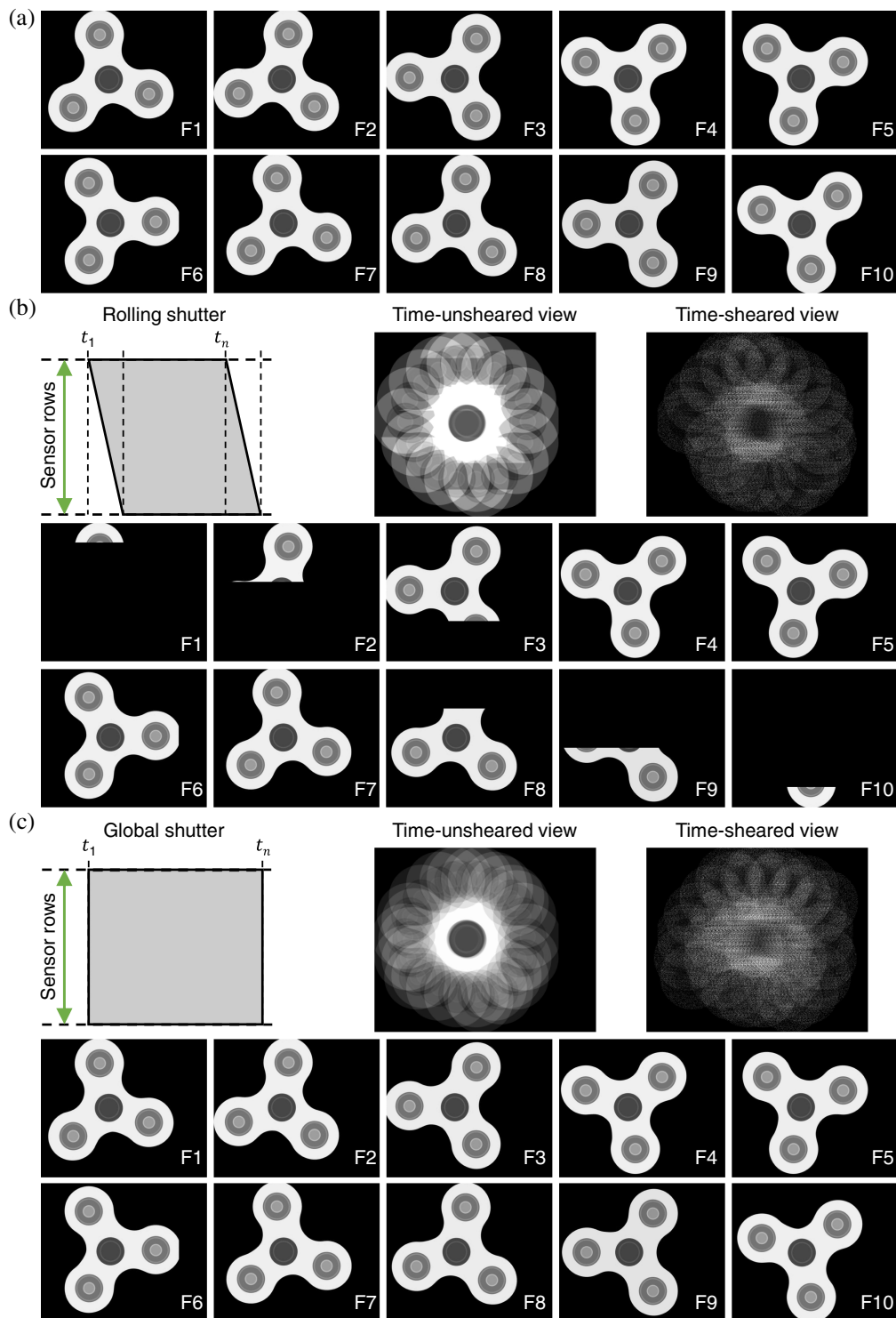
Fig. 10 Comparison between the rolling shutter and the global shutter in CUP's data acquisition. (a) 10 representative frames of a simulated rotating-spinner scene. (b) Rolling shutter's operating principle (top-left panel), the produced 2D snapshots (top-middle and top-right panels), and the illustration of CUP's reconstructed frames (bottom panel). (c) As (b), but showing the results produced by the global shutter. F1 to F10, frame indices.

point registration" in the "Image Processing Toolbox."[172] The function "cpselect" opens a window for the user to select at least four pairs of control points in both views. Then, the function "fitgeotform2d" estimates the transformation matrix that best aligns the control points. Finally, the function "imwarp" applies the transformation matrix to complete the co-registration. The
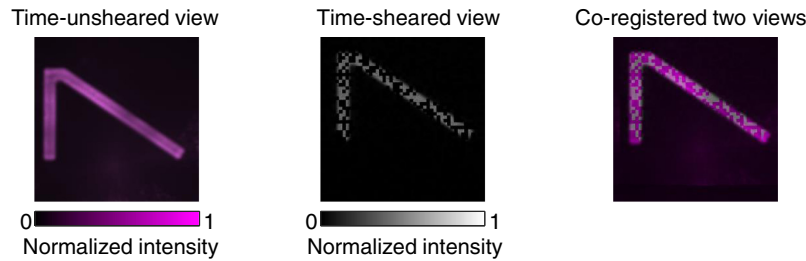
Time-unsheared view   Time-sheared view   Co-registered two views

0 — 1 Normalized intensity   0 — 1 Normalized intensity

Fig. 11 Co-registration for dual-view CUP.

co-registered time-unsheared view and the time-sheared view are then fed into the reconstruction algorithms.

3.4.2 *Acquisition of the encoding mask*

The experimentally captured encoding mask image produces better reconstruction than the design file of the used pattern because it takes into consideration various practical imperfections. For example, the DMD's micromirrors may have a different orientation than those of the sensors. The fixed encoders fabricated by high-definition printing or photolithography also have defective pixels or membrane curving. These cases cannot be eliminated even if the imaging system is tuned with a proper magnification that matches the size of the encoding pixels to that of the sensor's pixels. Thus, a mask image is captured by tuning off the shearing unit and then binarized for CUP's image reconstruction (Fig. 12). Besides background subtraction and white-field correction, threshold selection and edge detection are combined to optimize binarization.[57] This calibration can also reduce aberrations and field curvature.[123] In practice, the FOV and the maximum shearing distance are also limited to ensure high quality in captured images.

3.4.3 *Linearity test of shearing operation*

Linear temporal shearing is used in CUP's forward model (see Sec. 2.1). However, various experimental factors could deviate the linear temporal shearing operation, including misalignment, jitter, and imperfect instrument responses.[173] Therefore, a linearity test is required to assess the system's performance and to compensate for these factors. An example of a rotating-mirror-based CUP system is shown in Fig. 13(a). About 100 frames containing number indices and short lines were displayed on a DMD at 20 kHz. From the recorded snapshot, the displacements between the centroids of the consecutive short lines were calculated to determine the rotating
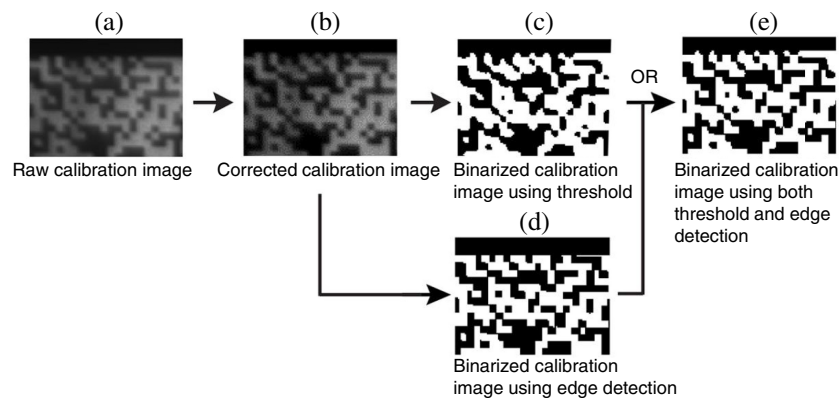

(a)   (b)   (c)   OR   (e)

Raw calibration image   Corrected calibration image   Binarized calibration image using threshold   Binarized calibration image using both threshold and edge detection

(d)

Binarized calibration image using edge detection

Fig. 12 Binarization of the captured encoding mask image. (a) Section cropped from the acquired mask image. (b) Cropped section after background subtraction and white-field correction. (c) Image binarization by applying a threshold to (b). (d) Image binarization by detecting edges in panel (b). (e) Combining (c) and (d) using OR operation. Adapted with permission from Ref. 57.
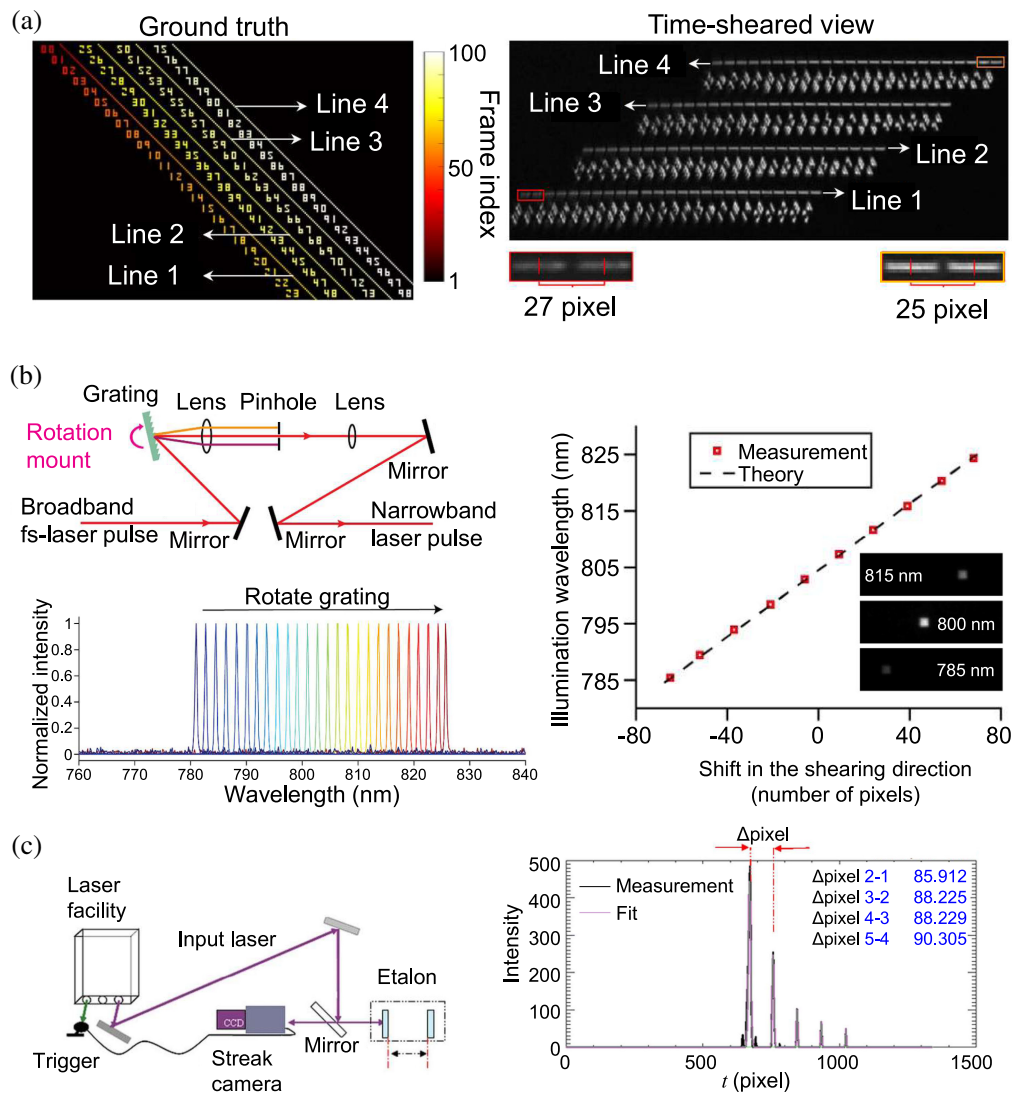
Fig. 13 Linearity test of the temporal shearing operation. (a) Test of a rotating-mirror-based temporal shearing. Left panel: composite of the 100 frames with consecutive short lines and frame indices. Right panel: analysis of displacement in the time-sheared view. (b) Test of a diffraction-grating-based temporal shearing. Left panel: setup (top) of generating pulses with selected wavelengths (bottom). Right panel: result of the linearity test by illuminating a square pattern with the narrow-band pulses. (c) Test of a streak-tube-based temporal shearing. Left panel: setup of the test. Right panel: cross-section in the streak measurement of the pulse train generated by the etalon. (a) Adapted with permission from Ref. 104. (b) Adapted with permission from Ref. 59. (c) Adapted with permission from Ref. 174.

mirror's shearing operation. In this example, the shearing deviates from a linear function by 2 pixels over 100 frames.[104]

As another example, Fig. 13(b) shows the linearity test of a diffraction-grating-based CUP system.[59] A tunable bandpass filter was built based on a rotating grating [top-left panel in Fig. 13(b)] to produce pulses with a selected wavelength [bottom-left panel in Fig. 13(b)]. The generated narrowband pulses illuminated a small square pattern, whose positions in the streak images were measured to obtain the relationship with wavelengths [right panel in Fig. 13(b)]. Finally, an example of the linearity test of an image-converter streak tube is shown in Fig. 13(c). Following a calibration protocol similar to that of the diffraction grating-based CUP system, a pulse train with a known interval was generated by an etalon. The linearity was computed by measuring the deflected pulses' positions.[174]

4 Biomedical Applications

Many biological processes, such as blood flow, brain activities, or cellular dynamics, are not repeatable. Single-shot CUP provides an innovative and complementary tool to probe these events, which generates valuable insights for the fundamental understanding of their underlying mechanisms. In this section, we will focus on four representative biomedical applications of CUP.

4.1 Neuroimaging

Monitoring the spatiotemporal dynamics of neuron signaling is essential to the understanding of the brain's structure and function. Direct visualization can aid researchers and clinicians in studying neurological disorders, cognitive processes, and brain development. Frame rates at the level of one billion fps are demanded to image the propagation of action potentials (APs) in myelinated axons (~100 m/s) with high spatial resolution and in real time. Unreachable by conventional electronic sensors, this requirement poses a considerable technical challenge to neuroimaging research.

Overcoming this challenge, CUP imaged phase and lifetime dynamics evoked by neuronal activities. As an example, by combining Mach–Zehnder interferometry and utilizing its ultrafast imaging speed and large sequence depth, differentially enhanced CUP (Diff-CUP) imaged internodal current flow in myelinated axons from the sciatic nerves of *Xenopus laevis* frog at 200×10^9 fps[65] [Fig. 14(a)]. The high phase sensitivity of Diff-CUP enables simultaneously capturing the substantial cellular deformations and consequent phase alterations induced by passive current flows (i.e., without the amplification of the electrical current)[175,176] resulting from a 10-V and 1-μs pulse injected into the axon [Fig. 14(b)]. The reconstructed correlation curves of each segment of the FOV (labeled with numbers 1 to 8) reveal the microsecond-level phase changes induced by the propagating internodal current flow [Fig. 14(c)], whose conduction speeds in myelinated axons were calculated to be 100 ± 26 m/s. To date, Diff-CUP is the fastest imaging-based approach for assessing AP-related conduction.

CUP is also implemented as a CS-based fluorescence lifetime imaging microscopy (FLIM)[60] to record high-resolution 2D lifetime images of immunofluorescently stained neurons [Fig. 14(d)]. With an imaging speed of $\sim 10 \times 10^9$ fps, this CUP-based FLIM system captured the fluorescence intensity decay in real time, which produced a 2D lifetime map. Leveraging the intrinsic frame rate of the internal CMOS camera, lifetime maps were generated at 100 fps. This technique visualized neural spike dynamics via the fluorescence intensity and donor lifetime decrease during Förster resonance energy transfer.[177] Figure 14(e) illustrates six representative lifetime images of a cultured hippocampal neuron at 100 fps. The time courses of the averaged fluorescence intensity variation and lifetime of this sample over 1 s are plotted in Fig. 14(f). Finally, the hippocampal neuron's fluorescence intensity and lifetime waveforms of single AP and their means [black lines and green line in Fig. 14(g), respectively] were acquired experimentally. This analysis revealed that a single spiking event led to an average relative fluorescence intensity change ($\Delta F/F$) of -2.9% and a lifetime change of -0.7 ns.

4.2 Temperature Sensing

Temperature, as an important biomarker, is linked to many biological processes (e.g., metabolism[178]) and medical procedures (e.g., photothermal therapy[179]). Accurate and real-time temperature sensing is important to pathology diagnostics, physiology monitoring, and therapeutical efficiency. Photoluminescence thermometry presents an emerging method by utilizing the temperature-sensitive optical emissions of photoluminescent materials as well as optical detections at high spatial resolution. Its merits include noncontactness, high adaptability to a broad temperature range, high accuracy, flexibility in sample selection, and suitability for diverse environments.[180] Thus, photoluminescence thermometry is increasingly featured in recent advances in optical temperature measurements.

The success of photoluminescence thermometry depends on two essential constituents: temperature indicators and optical imaging instruments. Recent advances in biochemistry, materials science, and molecular biology have unveiled numerous labeling indicators for photoluminescence thermometry.[162,181,182] From semiconductor quantum dots[183] and organic fluorophores[184]
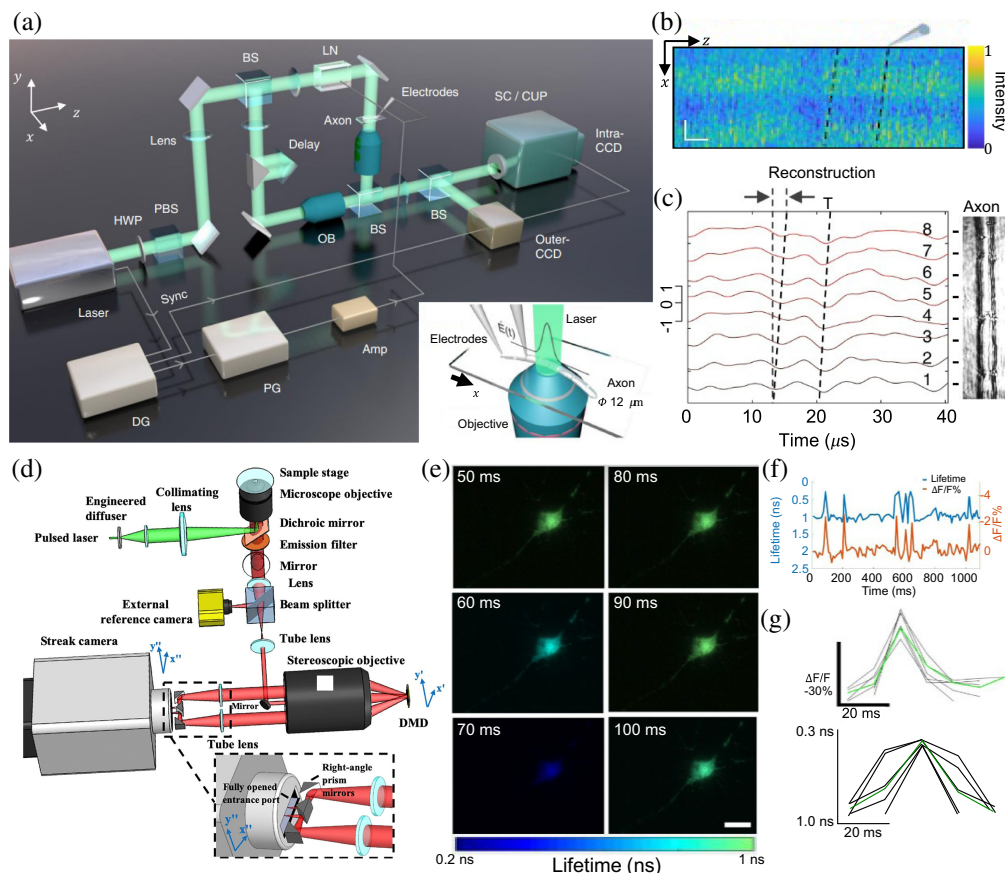
Fig. 14 CUP of neuronal activities. (a) Schematic of Diff-CUP. Inset: Adhesion microscope slide for Diff-CUP imaging. BS, beam splitter; DG, delay generator; $E(t)$, transient field stimulation; HWP, half-wave plate; LN, lithium niobate; OB, objective lens; PBS, polarizing beam splitter; PG, pulse generator; SC, streak camera. (b) Spatiotemporal interferogram of a propagating inter-nodal current flow in a myelinated axon captured by Diff-CUP. (c) Reconstruction of the current flow signals based on the stimulus interferogram. Black dashed lines indicate the signal region of the internodal current flow. T, the propagation time of the internodal current flow within the FOV. (d) Schematic of compressed FLIM. (e) Six representative frames from the reconstruction of a cultured hippocampal neuron upon potassium stimulation at 100 fps. (f) Time-lapsed lifetime and intensity curves of a cultured hippocampal neuron. (g) Intensity (top panel) and lifetime (bottom panel) waveforms of neural spikes (black lines) and their means (green lines) for a cultured hippo-campal neuron under stimulation. (a)–(c) Adapted with permission from Ref. 65. (d)–(g) Adapted with permission from Ref. 60.

to rare-earth-doped phosphors,[185] the diversity of these agents allows for tailored temperature sensing across different thermal sensitivities, optical properties, and response times for biomedi-cal applications.[186–189] For example, lanthanide-doped upconverting nanoparticles (UCNPs), which can sequentially absorb two (or more) low-energy near-infrared photons and convert them to one higher-energy photon, enable biocompatible temperature sensing with low excitation power densities and high sensitivity.[190,191]

CUP has enabled wide-field temperature mapping using photoluminescence lifetimes of UCNPs.[135] In the schematic shown in Fig. 15(a), near-infrared pulses, generated by a 980-nm continuous-wave (CW) laser and an optical chopper, are focused on the back focal plane of an objective lens to form wide-field illumination. The excited UCNPs on the sample emit visible upconversion luminescence. After passing the filter, the dynamic photoluminescence of a selected emission band is imaged by a rotating mirror-based dual-view CUP system at 33,000 fps. The reconstructed lifetime images in the UCNPs' two upconversion emission bands at different temperatures are shown in Fig. 15(b). The averaged intensity decays [Fig. 15(c)]
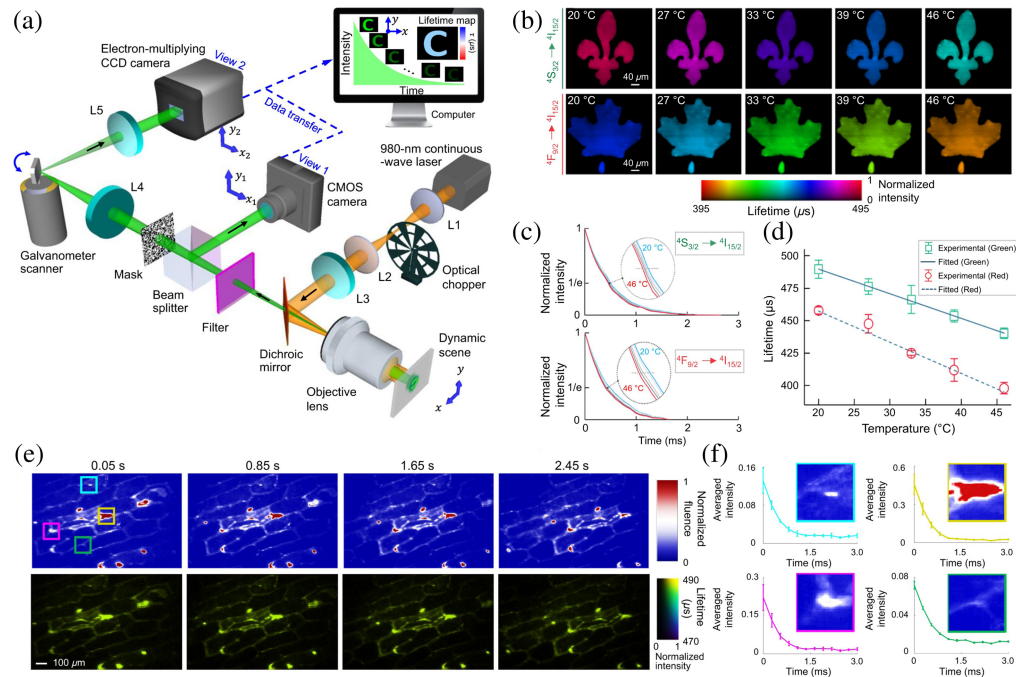
Fig. 15 CUP of temperature sensing. (a) Schematic of wide-field photoluminescence lifetime thermometry based on a dual-view rotating-mirror CUP system. L1 to L5, lenses. (b) Lifetime maps of the two emission bands (i.e., $^4S_{3/2} \rightarrow {}^4I_{15/2}$ and $^4F_{9/2} \rightarrow {}^4I_{15/2}$) of the used UCNPs under different temperatures. (c) Normalized photoluminescence decays of the two emission bands after averaging over the FOV. (d) Temperature-lifetime relationship of both emission bands. (e) Selected time-unsheared views (top row) and reconstructed lifetime maps (bottom row) of a moving onion epidermis cell sample labeled by UCNPs. (f) Intensity decays at four selected areas with different intensities marked in (e). Adapted with permission from Ref. 135.

enable the establishment of the temperature-lifetime relationship [Fig. 15(d)]. Furthermore, the system tracked the 2D temperature of a moving onion epidermis sample labeled by UCNPs at a rate of 20 lifetime-maps per second [Fig. 15(e)]. The intensity decays of four selected areas [labeled in the top-left panel in Fig. 15(e)] are shown in Fig. 15(f). It is worth noting that the fluences of the four selected areas are different but the measured photoluminescence lifetimes remain stable, showing the lifetime-based approach contributed by CUP is more reliable in accurate temperature sensing.

4.3 Microfluidics

A rotating-mirror CUP system has been applied to the video recording of complex fluid dynamics and interactions at the microscale.[148] A schematic of rotating-mirror-based CUP is shown in Fig. 16(a). This system observed flow droplet samples within a microfluidic chip.[149] Two immiscible liquids of transparent oil and chemical dye, injected through a motorized dispenser, flowed in the chip channels at 0.9 m/s. Three separate measurements are recorded at 3000 fps, 50,000 fps, and 120,000 fps [Fig. 16(b)]. These experimental results show a high reconstruction quality with well-preserved edge features in the frames, showing clear and distinguishable droplets flowing in the microfluidic chip. These results show CUP's potential to visualize cell-shape changes in response to rapid external stimuli or internal dynamics in microfluidics,[192–194] which will provide new insights into cellular biomechanical properties that are closely linked to cellular function and disease development.

4.4 Photoacoustic Imaging

CUP can also contribute to photoacoustic (PA) imaging. Figure 17 shows a simulation study on implementing CUP with optical interferometric detection of PA waves.[195,196] In the proposed system schematically shown in Fig. 17(a), a pulsed laser illuminates a biological sample.
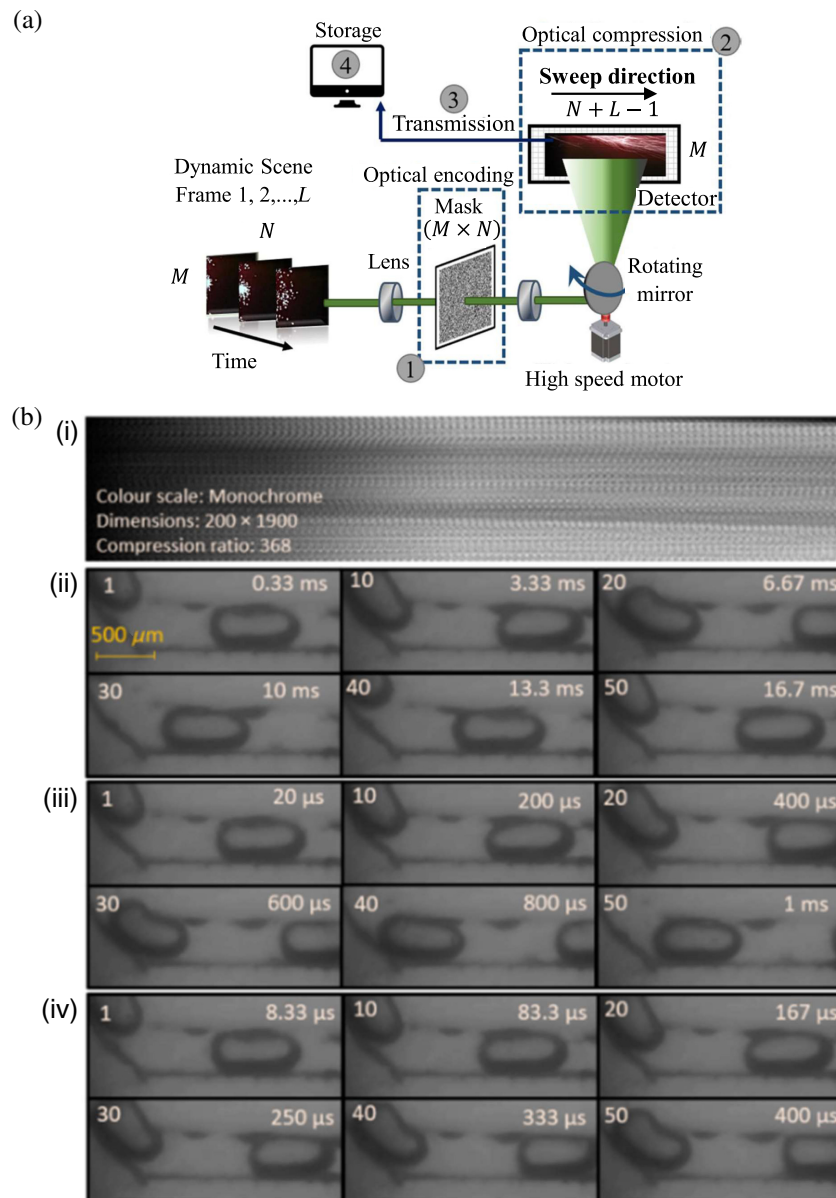
Fig. 16 CUP of microfluidics. (a) Schematic of a rotating-mirror-based CUP system. (b) Snapshot of flowing immiscible liquids (i) and representative frames from the reconstructed videos at (ii) 3000, (iii) 50,000, and (iv) 120,000 fps. Adapted with permission from Ref. 148 and Ref. 149.

The induced PA effect generates thermoelastic initial pressure, which is detected at the surface of the sample with a Fabry–Pérot etalon (FPE). This interaction of the ultrasonic waves with the surface of the FPE results in the modulation of the reflected CW laser beam on the opposite side of the FPE.[197] The modulated CW laser beam is then imaged by a CUP system based on a DMD and a galvanometer scanner. Figure 17(b) shows a simulation of this method to image the initial pressure distribution of 12 vessel-like structures.

5 Prospect

CUP has largely advanced ultrafast imaging instrumentation. Its generic sensing model indicates that deployed components, rather than the theory, limit the system's performance. Therefore, CUP has vast potential to be further improved in its imaging capability. In this section, we outline seven aspects of CUP's future technical development.
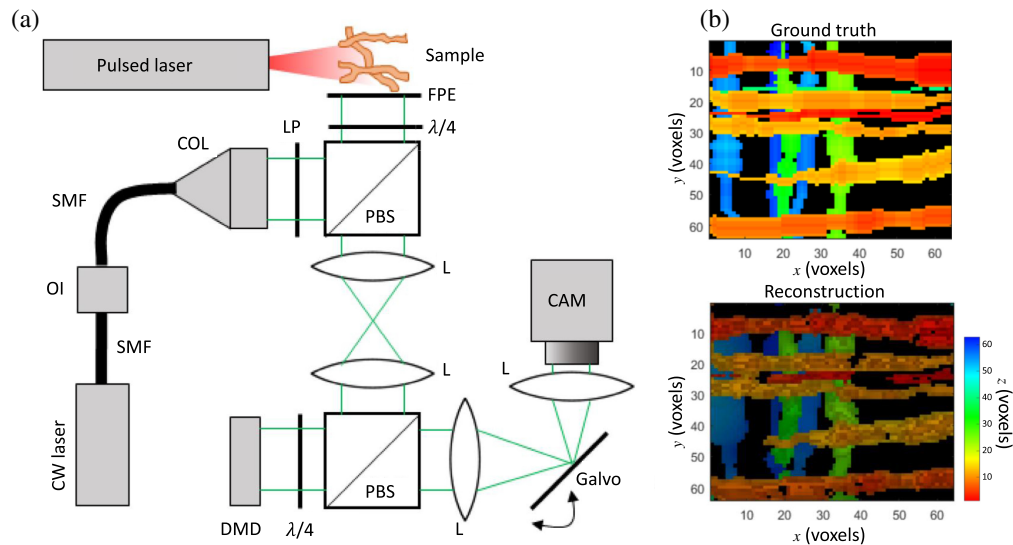
Fig. 17 CUP of photoacoustic imaging. (a) Proposed system schematic. CAM, camera; COL, collimator; CW, continuous-wave; DMD, digital micromirror device; FPE, Fabry–Pérot etalon; L, lens; $\lambda/4$, quarter wave plate; LP, linear polarizer; OI, optical isolator; PBS, polarizing beam splitter; SMF, single-mode fiber. (b) Simulation of image reconstruction of initial pressure distribution. Adapted with permission from Ref. 195.

5.1 Faster

As currently the world's fastest optical imaging technology, CUP naturally carries a mission to explore even higher speeds in optical imaging. Since CUP's invention, innovation in temporal shearing units has been a focus of its technical improvement. From 100×10^9 fps of the original CUP system,[17] various image-converter streak tubes have been deployed to increase its imaging speed to 10×10^{12} fps,[114] which currently holds the world record for single-shot receive-only ultrafast imaging. However, at this speed, the image quality is considerably affected by the space-charge effect,[160] posing challenges in further improvement of frame rates. In the future, transient perturbation in reflective index induced by a temporally modulated ultrashort laser pulse or molecular orientation could bring higher imaging speeds and circumvent image degradation.[151,170]

Leveraging the advance in chirped pulse illumination, CUP systems using passive temporal shearing units have boosted the imaging speed to 3.85×10^{12} fps,[121] 70×10^{12} fps,[57] 219×10^{12} fps,[198] and 256×10^{12} fps.[63] The last value marks the fastest speed in single-shot optical imaging. In the future, by synergizing the ultra-broadband ultrashort pulses[199] and photonic streaking in gas,[170,200] CUP's imaging speed could top quadrillion fps, entering the attosecond-level imaging regime.

5.2 Clearer

A higher spatial resolution allows CUP to visualize fine details. In biomedicine, this ability can transfer to informative depiction of cellular and tissue morphology, accurate diagnostics, and precise treatment. Nonetheless, in CUP's operation, both spatial encoding in data acquisition and denoising in image reconstruction could reduce the effective system bandwidth. To visualize the targeted spatial details, a common practice is to magnify the scene, which unavoidably reduces the FOV. Thus, how to regain the lost bandwidth to achieve a diffraction-limited spatial resolution is an important research direction of CUP. One potential approach is subpixel shifting.[201] In particular, a DOE could be used to duplicate the dynamic scene to multiple bands, each would be encoded with the same encoding mask but with a different subpixel shift. A joint image reconstruction using all the captured snapshots could recover the original optical bandwidth.

Another interesting research direction is super-resolution CUP. As many ultrafast phenomena also occur at the nanoscale, overcoming the diffraction limit in the CUP system will likely open avenues for new studies not possible before, including temperature dynamics in mitochondria,[202] conformational transitions of protein,[203] evolutions of membrane fragments produced by

cellular lysis.[204] Toward this goal, CUP can be incorporated into existing super-resolution microscopy techniques (e.g., structured illumination microscopy[205]) or bypass the optical diffraction limit by electron imaging (e.g., transmission electron microscopy[117]).

5.3 Broader Spectrum

Although CUP has been experimentally demonstrated in the ultraviolet, visible, and near-infrared spectral ranges, extending its imaging capability to a broader spectrum will likely continue in future application-driven development. Toward this goal, the spatial encoder should have high contrast in the desired spectrum. Besides the popular broadband metallic masks made from aluminum, silver, or chromium,[17,91,147] photonic crystals with broad tunable bandgaps can selectively block specific wavelengths,[206] giving them the potential to be used as a spatial encoder in certain spectra. For temporal shearing units, the photocathode in the image-converter streak tube excludes high-sensitivity imaging for wavelengths of >950 nm. Contrarily, leveraging its all-optical functionality, rotating mirrors can fill out this gap, which will likely lead to the development of CUP for deep ultraviolet, mid-infrared, and far-infrared spectra. Moreover, advanced design and fabrication of metasurfaces and metalenses could potentially extend CUP to a spectrum from extreme ultraviolet to terahertz.[207,208]

5.4 Smarter

Many deep learning-based approaches have been used in CUP's image reconstruction.[94–104] Harnessing the power of artificial intelligence, they have unlocked new capabilities for analyzing ultrafast events, such as real-time data processing, on-device analysis, and on-time feedback. It is expected that these deep-learning algorithms will provide multifacet supervision to CUP systems in the future. For example, the next-generation systems could autonomously adjust the patterns loaded on the spatial encoder according to the initial classification of the dynamic scene. These systems could also monitor the nonlinear shearing operation and adaptively compensate for it in image reconstruction or system alignment.[104]

5.5 Higher Dimensions

Recent developments in CUP have explored high-dimensional ultrafast imaging. To date, several advanced systems—such as multispectral CUP,[120] stereo-polarimetric CUP,[57] and spectral-volumetric CUP[62]—have pushed the overall sensing capability to four dimensions and even five dimensions. In the future, by extending the configuration used in stereo-polarimetric CUP[57] to generate multiple perspectives of the dynamic scene, light-field imaging could be incorporated into CUP. Ultimately, single-shot imaging of seven-dimensional plenoptic function would be within reach. Using CUP to sense other photon tags that are not included in the conventionally defined plenoptic function is also a future direction. CUP has already enabled amplitude and phase imaging of a femtosecond laser pulse.[63] CUP could also be combined with other existing technologies, such as the transport-of-intensity equation [209] and coherent modulation,[210] for ultrafast quantitative phase imaging. Finally, recent advancements in on-chip polarization imaging and metasurface-based angular moment separation could incorporate these parameters into CUP's measurement scope.[211,212]

5.6 Smaller

CUP systems with a compact size are important to studies that require restricted weight and space. For biomedicine, it will offer the ability to mount the system the same way as conventional cameras on microscopes and hand-held systems as well as in operating rooms. An innovative optical design that folds the optical path could reduce the system size.[213] Selecting a multifunctional component (e.g., a metalens and a TDI sensor) provides another approach to reducing the number of optical elements in CUP systems. Advances in sensor design and nanofabrication could provide the streak imaging sensor[167,168] with a 2D FOV. All efforts will contribute to engineering compact and even miniature CUP systems in the future.

5.7 Cheaper

Besides reducing the size of a CUP system, making an economical CUP system carries considerable value from both research and commercialization perspectives. To manufacture a fixed

spatial encoder, high-resolution printing offers the lowest cost (i.e., <\$20 per in.^{2}). For reconfigurable spatial encoders, a 0.47″ DMD chip (1920 × 1080 micromirrors; 5.4-μm pitch) costs ~\$120.[214] Future development could develop a DMD controller specifically tailored for CUP with much-reduced functionality compared with existing ones to decrease the cost. For the temporal shearing unit, the first approach to reduce the cost is to find a replacement for expensive image-converter streak tubes. Electro-optic modulators have made their debut in this direction,[122] which produced an imaging speed of 50×10^9 fps. For a rotating-mirror CUP system, a viable strategy is to add the spatial encoder and an affordable rotating mirror (e.g., galvanometer scanners and polygonal mirrors) in front of the CCD/CMOS cameras existing in the system. It is envisaged that a minor addition in cost could endow ultrahigh-speed imaging to existing cameras while retaining their inherent advantages (e.g., in sensitivity and sensing spectrum).

6 Conclusions

In this tutorial, we have elucidated the fundamentals of CUP. We have provided Matlab codes that create CUP's sensing matrices and simulate the acquired snapshots based on the forward model. Matlab/Python codes and examples are also included for two respective reconstruction algorithms—one based on analytical modeling using the ADMM and the other on deep learning using the D-HAN. To facilitate comprehension, a "cell-division" scene is simulated step by step alongside the provided codes. A fully operational CUP system relies on three essential hardware components: a spatial encoder, a temporal shearing unit, and a 2D sensor. We have surveyed representative implementations of each component as well as calibration steps in both data acquisition and image reconstruction.

Ever since its invention, CUP has stayed under the spotlight in research as an emerging and innovative imaging platform. Its evolution has been shaped by the innovation of imaging strategies and the adaptive optimization of its key components, leading to its widespread implementation in various biomedical applications. CUP—as currently the world's fastest single-shot optical imaging modality—is positioned for future advancements in imaging speed, spatial resolution, sensing spectrum, artificial intelligence, imaging dimensionality, system size, and manufacturing cost. CUP is highly anticipated to make more remarkable progress in biomedicine.

References

1. J. Liang and L. V. Wang, "Ultrafast optical imaging," in *Handbook of Laser Technology Applications*, C. Guo and S. C. Singh, Eds., pp. 315–328, CRC Press (2021).
2. H. Roder, "Stepwise helix formation and chain compaction during protein folding," *Proc. Natl. Acad. Sci.* **101**(7), 1793–1794 (2004).

3. D. N. Ku, "Blood flow in arteries," *Annu. Rev. Fluid Mech.* **29**(1), 399–434 (1997).

4. C. A. Day and A. K. Kenworthy, "Tracking microdomain dynamics in cell membranes," *Biochim. Biophys. Acta Biomembr.* **1788**(1), 245–253 (2009).

5. H. Astacio, A. Vasin, and M. Bykhovskaia, "Stochastic properties of spontaneous synaptic transmission at individual active zones," *J. Neurosci.* **42**(6), 1001–1019 (2022).

6. F. Vetrone et al., "Temperature sensing using fluorescent nanothermometers," *ACS Nano* **4**(6), 3254–3258 (2010).

7. Y. Liu et al., "Optical focusing deep inside dynamic scattering media with near-infrared time-reversed ultrasonically encoded (TRUE) light," *Nat. Commun.* **6**(1), 5904 (2015).

8. J. Liang and L. V. Wang, "Single-shot ultrafast optical imaging," *Optica* **5**(9), 1113–1127 (2018).

9. H. Mikami, L. Gao, and K. Goda, "Ultrafast optical imaging technology: principles and applications of emerging methods," *Nanophotonics* **5**, 497–509 (2016).

10. Q. Miao et al., "Molecular afterglow imaging with bright, biodegradable polymer nanoparticles," *Nat. Biotechnol.* **35**(11), 1102–1110 (2017).

11. P. S. May and M. Berry, "Tutorial on the acquisition, analysis, and interpretation of upconversion luminescence data," *Method. Appl. Fluoresc.* **7**(2), 023001 (2019).

12. R. Kodama et al., "Fast heating of ultrahigh-density plasma as a step towards laser fusion ignition," *Nature* **412**(6849), 798–802 (2001).

13. H. Zhang et al., "Efficient optical Kerr gate of tellurite glass for acquiring ultrafast fluorescence," *J. Opt.* **14**(6), 065201 (2012).

14. J. Wang et al., "Beam parameters measurement with a streak camera in HLS," in *Proc. of PAC09 (2009)*.

15. T. G. Etoh et al., "A 16 Mfps 165 kpixel backside-illuminated CCD," in *IEEE Int. Solid-State Circuits Conf.*, pp. 406–408 (2011).

16. D. Qi et al., "Single-shot compressed ultrafast photography: a review," *Adv. Photonics* **2**(1), 014003 (2020).

17. L. Gao et al., "Single-shot compressed ultrafast photography at one hundred billion frames per second," *Nature* **516**(7529), 74–77 (2014).

18. J. Liang, "Punching holes in light: recent progress in single-shot coded-aperture optical imaging," *Rep. Prog. Phys.* **83**(11), 116101 (2020).

19. J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Adv. Opt. Photonics* **10**(2), 409–483 (2018).

20. J. Hunt et al., "Metamaterial apertures for computational imaging," *Science* **339**(6117), 310–313 (2013).

21. W. Oh et al., "High-speed polarization sensitive optical frequency domain imaging with frequency multiplexing," *Opt. Express* **16**(2), 1096–1103 (2008).

22. A. Ehn et al., "FRAME: femtosecond videography for atomic and molecular dynamics," *Light Sci. Appl.* **6**(9), e17045 (2017).

23. H. Li et al., "Investigation of single-shot high-speed photography based on spatial frequency multiplexing," *J. Opt. Soc. Am. A* **40**(3), 521–529 (2023).

24. Q. Yue et al., "One-shot time-resolved holographic polarization microscopy for imaging laser-induced ultrafast phenomena," *Opt. Express* **25**(13), 14182–14191 (2017).

25. K. Nakagawa et al., "Sequentially timed all-optical mapping photography (STAMP)," *Nat. Photonics* **8**(9), 695–700 (2014).

26. T. Suzuki et al., "Sequentially timed all-optical mapping photography (STAMP) utilizing spectral filtering," *Opt. Express* **23**(23), 30512–30522 (2015).

27. M. Touil et al., "Acousto-optically driven lensless single-shot ultrafast optical imaging," *Light Sci. Appl.* **11**(1), 66 (2022).

28. G. Gao et al., "Ultrafast all-optical solid-state framing camera with picosecond temporal resolution," *Opt. Express* **25**(8), 8721–8729 (2017).

29. X. Zeng et al., "High-spatial-resolution ultrafast framing imaging at 15 trillion frames per second by optical parametric amplification," *Adv. Photonics* **2**(5), 056002 (2020).

30. X. Zeng et al., "Review and prospect of single-shot ultrafast optical imaging by active detection," *Ultrafast Sci.* **3**, 0020 (2023).

31. S. S. Harilal et al., "Plume splitting and sharpening in laser-produced aluminium plasma," *J. Phys. D Appl. Phys.* **35**(22), 2935 (2002).

32. Y. Fang et al., "A four-channel ICCD framing camera with nanosecond temporal resolution and high spatial resolution," *J. Mod. Opt.* **68**(13), 661–669 (2021).

33. L. Cester et al., "Time-of-flight imaging at 10 ps resolution with an ICCD camera," *Sensors* **19**(1), 180 (2019).

34. X. Liu et al., "Diffraction-gated real-time ultrahigh-speed mapping photography," *Optica* **10**(9), 1223–1230 (2023).

35. Y. Tsuchiya and Y. Shinoda, "Recent developments of streak cameras," *Proc. SPIE* **0533**, 110–116 (1985).

36. T. G. Etoh et al., "The theoretical highest frame rate of silicon image sensors," *Sensors* **17**(3), 483 (2017).

37. X. Liu et al., "Single-shot compressed optical-streaking ultra-high-speed photography," *Opt. Lett.* **44**(6), 1387–1390 (2019).

38. R. M. Willett, R. F. Marcia, and J. M. Nichols, "Compressed sensing for practical optical imaging systems: a tutorial," *Opt. Eng.* **50**(7), 072601 (2011).

39. C. Zuo et al., "Deep learning in optical metrology: a review," *Light Sci. Appl.* **11**(1), 39 (2022).

40. M. A. Alonso, "Wigner functions in optics: describing beams as ray bundles and pulses as particle ensembles," *Adv. Opt. Photonics* **3**(4), 272–365 (2011).

41. N. Hagen et al., "Snapshot advantage: a review of the light collection improvement for parallel high-dimensional measurement systems," *Opt. Eng.* **51**(11), 111702 (2012).

42. J. Yao et al., "Exploring femtosecond laser ablation by snapshot ultrafast imaging and molecular dynamics simulation," *Ultrafast Sci.* **2022**, 9754131 (2022).

43. P. Wang and L. V. Wang, "Compressed ultrafast photography," in *Coded Optical Imaging*, J. Liang, Ed., Springer Nature (2023).

44. D. Faccio and A. Velten, "A trillion frames per second: the techniques and applications of light-in-flight photography," *Rep. Prog. Phys.* **81**(10), 105901 (2018).

45. M. GarciaLechuga, J. Solis, and J. Siegel, "Probing matter by light," in *Ultrafast Laser Nanostructuring: The Pursuit of Extreme Scales*, R. Stoian and J. Bonse, Eds., pp. 277–319, Springer International Publishing, Cham (2023).

46. F. S. Oktem, L. Gao, and F. Kamalabadi, "Computational spectral and ultrafast imaging via convex optimization," in *Handbook of Convex Optimization Methods in Imaging Science*, V. Monga, Ed., pp. 105–127, Springer International Publishing, Cham (2018).

47. L. Gao and L. V. Wang, "A review of snapshot multidimensional optical imaging: measuring photon tags in parallel," *Phys. Rep.* **616**, 1–37 (2016).

48. K. Uchiyama et al., "Various ultra-high-speed imaging and applications by Streak camera," in *3rd Int. Workshop Image Sensor and Systems, invited-12* (2016).

49. P. Llull et al., "Coded aperture compressive temporal imaging," *Opt. Express* **21**(9), 10526–10545 (2013).

50. Y. Sun, X. Yuan, and S. Pang, "Compressive high-speed stereo imaging," *Opt. Express* **25**(15), 18182–18190 (2017).

51. M. Qiao, X. Liu, and X. Yuan, "Snapshot spatial–temporal compressive imaging," *Opt. Lett.* **45**(7), 1659–1662 (2020).

52. X. Yuan, Y. Sun, and S. Pang, "Compressive video sensing with side information," *Appl. Opt.* **56**(10), 2697–2704 (2017).

53. L. Wang et al., "Spatial-temporal transformer for video snapshot compressive imaging," *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(7), 9072–9089 (2023).

54. X. Yuan et al., "Plug-and-play algorithms for video snapshot compressive imaging," *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(10), 7093–7111 (2022).

55. L. Zhu et al., "Space- and intensity-constrained reconstruction for compressed ultrafast photography," *Optica* **3**(7), 694–697 (2016).

56. H. Gao et al., "A simple yet effective AIE-based fluorescent nano-thermometer for temperature mapping in living cells using fluorescence lifetime imaging microscopy," *Nanoscale Horiz.* **5**(3), 488–494 (2020).

57. J. Liang et al., "Single-shot stereo-polarimetric compressed ultrafast photography for light-speed observation of high-dimensional optical transients with picosecond resolution," *Nat. Commun.* **11**(1), 5252 (2020).

58. J. Liang et al., "Single-shot real-time video recording of a photonic Mach cone induced by a scattered light pulse," *Sci. Adv.* **3**(1), e1601814 (2017).

59. P. Wang, J. Liang, and L. V. Wang, "Single-shot ultrafast imaging attaining 70 trillion frames per second," *Nat. Commun.* **11**(1), 2091 (2020).

60. Y. Ma et al., "High-speed compressed-sensing fluorescence lifetime imaging microscopy of live cells," *Proc. Natl. Acad. Sci.* **118**(3), e2004176118 (2021).

61. T. Kim et al., "Picosecond-resolution phase-sensitive imaging of transparent objects in a single shot," *Sci. Adv.* **6**(3), eaay6200 (2020).

62. P. Ding et al., "Single-shot spectral-volumetric compressed ultrafast photography," *Adv. Photonics* **3**(4), 045001 (2021).

63. H. Tang et al., "Single-shot compressed optical field topography," *Light Sci. Appl.* **11**(1), 244 (2022).

64. J. Liang et al., "Encrypted three-dimensional dynamic imaging using snapshot time-of-flight compressed ultrafast photography," *Sci. Rep.* **5**(1), 15504 (2015).

65. Y. Zhang et al., "Ultrafast and hypersensitive phase imaging of propagating internodal current flows in myelinated axons and electromagnetic pulses in dielectrics," *Nat. Commun.* **13**(1), 5247 (2022).

66. C. Yang et al., "Compressed ultrafast photography by multi-encoding imaging," *Laser Phys. Lett.* **15**(11), 116202 (2018).

67. M. Cicconet et al., "Label free cell-tracking and division detection based on 2D time-lapse images for lineage analysis of early embryo development," *Comput. Biol. Med.* **51**, 24–34 (2014).
68. M. Marquez, Y. Lai, and J. Liang, "CUP-Tutorial_Data_1," (2023). https://drive.google.com/file/d/1aAD4DfSimEth25aSoDQkVIpNQvPjPXvh/view?usp=drive_link.
69. S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: fixed-point convergence and applications," *IEEE Trans. Comput. Imaging* **3**(1), 84–98 (2016).
70. S. Sreehari et al., "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Trans. Comput. Imaging* **2**(4), 408–423 (2016).
71. A. Gnanasambandam et al., "Megapixel photon-counting color imaging using quanta image sensor," *Opt. Express* **27**(12), 17298–17310 (2019).
72. Y. Sun et al., "Scalable plug-and-play ADMM with convergence guarantees," *IEEE Trans. Comput. Imaging* **7**, 849–863 (2021).
73. A. M. Teodoro, J. M. Bioucas-Dias, and M. A. Figueiredo, "A convergent image fusion algorithm using scene-adapted Gaussian-mixture-based denoising," *IEEE Trans. Image Process.* **28**(1), 451–463 (2018).
74. W. Dong et al., "Model-guided deep hyperspectral image super-resolution," *IEEE Trans. Image Process.* **30**, 5754–5768 (2021).
75. M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signal Process.* **1**(4), 586–597 (2007).
76. S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.* **57**(7), 2479–2493 (2009).
77. M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.* **20**(3), 681–695 (2010).
78. J. M. Bioucas-Dias and M. A. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Image Process.* **16**(12), 2992–3004 (2007).
79. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D* **60**(1), 259–268 (1992).
80. M. Marquez et al., "Snapshot compressive spectral depth imaging from coded aberrations," *Opt. Express* **29**(6), 8142–8159 (2021).
81. M. Marquez et al., "Compressive spectral imaging via deformable mirror and colored-mosaic detector," *Opt. Express* **27**(13), 17795–17808 (2019).
82. M. Marquez, H. Rueda-Chacon, and H. Arguello, "Compressive spectral light field image reconstruction via online tensor representation," *IEEE Trans. Image Process.* **29**, 3558–3568 (2020).
83. X. Yuan et al., "Plug-and-play algorithms for large-scale snapshot compressive imaging," in *Proc. IEEE/CVF Conf. Comput. Vis. and Pattern Recognit.*, pp. 1447–1457 (2020).
84. S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**(1), 1–122 (2011).
85. D. Knowles, "Lagrangian duality for dummies," (2010). https://www-cs.stanford.edu/~davidknowles/lagrangian_duality.pdf (accessed 23 December 2023).
86. S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).
87. N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.* **1**(3), 127–239 (2014).
88. S. H. Chan, "Performance analysis of plug-and-play ADMM: a graph signal processing perspective," *IEEE Trans. Comput. Imaging* **5**(2), 274–286 (2019).
89. E. T. Reehorst and P. Schniter, "Regularization by denoising: clarifications and new interpretations," *IEEE Trans. Comput. Imaging* **5**(1), 52–67 (2018).
90. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
91. Y. Lai et al., "Single-shot ultraviolet compressed ultrafast photography," *Laser Photonics Rev.* **14**(10), 2000122 (2020).
92. M. Marquez, Y. Lai, and J. Liang, "CUP-Tutorial_Data_2," (2023). https://drive.google.com/drive/folders/1lc6W9zg79T2dHSjyxC0oYTt8nuyCNeqS?usp=drive_link (accessed 23 December 2023).
93. I. Lemhadri et al., "LassoNet: a neural network with feature sparsity," *J. Mach. Learn. Res.* **22**(127), 5633–5661 (2021).
94. M. Qiao et al., "Deep learning for video compressive sensing," *APL Photonics* **5**(3), 030801 (2020).
95. A. Zhang et al., "Single-shot compressed ultrafast photography based on U-net network," *Opt. Express* **28**(26), 39299–39310 (2020).
96. C. Yang et al., "High-fidelity image reconstruction for compressed ultrafast photography via an augmented-Lagrangian and deep-learning hybrid algorithm," *Photonics Res.* **9**(2), B30–B37 (2021).
97. X. Liu et al., "Single-shot real-time compressed ultrahigh-speed imaging enabled by a snapshot-to-video autoencoder," *Photonics Res.* **9**(12), 2464–2474 (2021).

98. Q. Shen, J. Tian, and C. Pei, "A novel reconstruction algorithm with high performance for compressed ultrafast imaging," *Sensors* **22**(19), 7372 (2022).

99. C. Yang et al., "Improving the image reconstruction quality of compressed ultrafast photography via an augmented Lagrangian algorithm," *J. Opt.* **21**(3), 035703 (2019).

100. Z. Kaitao et al., "CUP-VISAR image reconstruction based on low-rank prior and total-variation regularization," *High Power Laser Part. Beams* **35**(7), 072002 (2023).

101. X. Wang et al., "Research of CUP-VISAR velocity reconstruction based on weighted DRUNet and total variation joint optimization," *Opt. Lett.* **48**(20), 5181–5184 (2023).

102. Y. Ma, X. Feng, and L. Gao, "Deep-learning-based image reconstruction for compressed ultrafast photography," *Opt. Lett.* **45**(16), 4400–4403 (2020).

103. Y. He et al., "High-speed compressive wide-field fluorescence microscopy with an alternant deep denoisers-based image reconstruction algorithm," *Opt. Lasers Eng.* **165**, 107541 (2023).

104. M. Marquez et al., "Deep-learning supervised snapshot compressive imaging enabled by an end-to-end adaptive neural network," *IEEE J. Sel. Top. Signal Process.* **16**(4), 688–699 (2022).

105. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).

106. W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.* **31**(2), 221–239 (1989).

107. H. Zhao et al., "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imaging* **3**(1), 47–57 (2016).

108. M. Gygli et al., "Creating summaries from user videos," *Lect. Notes Comput. Sci.* **8695**, 505–520 (2014).

109. H. Kiani Galoogahi et al., "Need for speed: a benchmark for higher frame rate object tracking," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 1125–1134 (2017).

110. S. M. Safdarnejad et al., "Sports videos in the wild (SVW): a video dataset for sports analysis," in *11th IEEE Int. Conf. and Workshops on Autom. Face and Gesture Recognit. (FG)*, pp. 1–7 (2015).

111. M. Marquez, Y. Lai, and J. Liang, "CUP-Tutorial_Data_3," (2023). https://drive.google.com/drive/folders/1-RL-8EKG-evQ1CJoQD9GS4N_BzWGTBj9?usp=drive_link (accessed 23 December 2023).

112. M. Marquez, Y. Lai, and J. Liang, "CUP-Tutorial_Data_4," (2023). https://drive.google.com/drive/folders/1Bt1yldG9Bik3C8I56a8nOxyyzlHkWIMu?usp=drive_link (accessed 23 December 2023).

113. W. Zhou et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).

114. J. Liang, L. Zhu, and L. V. Wang, "Single-shot real-time femtosecond imaging of temporal focusing," *Light Sci. Appl.* **7**(1), 42 (2018).

115. E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory* **52**(12), 5406–5425 (2006).

116. C. Yang et al., "Optimizing codes for compressed ultrafast photography by the genetic algorithm," *Optica* **5**(2), 147–151 (2018).

117. X. Liu et al., "Single-shot real-time sub-nanosecond electron imaging aided by compressed sensing: analytical modeling and simulation," *Micron* **117**, 47–54 (2019).

118. M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, "Deep fully-connected networks for video compressive sensing," *Digit. Signal Process.* **72**, 9–18 (2018).

119. M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, "DeepBinaryMask: learning a binary mask for video compressive sensing," *Digit. Signal Process.* **96**, 102591 (2020).

120. C. Yang et al., "Hyperspectrally compressed ultrafast photography," *Phys. Rev. Lett.* **124**(2), 023902 (2020).

121. Y. Lu et al., "Compressed ultrafast spectral-temporal photography," *Phys. Rev. Lett.* **122**(19), 193904 (2019).

122. C. Yang et al., "Single-shot receive-only ultrafast electro-optical deflection imaging," *Phys. Rev. Appl.* **13**(2), 024001 (2020).

123. Y. N. Mishra et al., "Single-pulse real-time billion-frames-per-second planar imaging of ultrafast nanoparticle-laser dynamics and temperature in flames," *Light Sci. Appl.* **12**(1), 47 (2023).

124. F. Cao et al., "Single-shot spatiotemporal intensity measurement of picosecond laser pulses with compressed ultrafast photography," *Opt. Lasers Eng.* **116**, 89–93 (2019).

125. C. Jin et al., "Weighted multi-scale denoising via adaptive multi-channel fusion for compressed ultrafast photography," *Opt. Express* **30**(17), 31157–31170 (2022).

126. S. Nagai et al., "Construction of multi-directional high-speed imaging system using a streak camera and image compression method," *High Energy Density Phys.* **37**, 100902 (2020).

127. Z. Guan et al., "Study on the length of diagnostic time window of CUP-VISAR," *Meas. Sci. Technol.* **32**(12), 125208 (2021).

128. C. Jin et al., "Single-shot real-time imaging of ultrafast light springs," *Sci. China Phys. Mech. Astron.* **64**(12), 124212 (2021).

129. L. Fan et al., "Real-time observation and control of optical chaos," *Sci. Adv.* **7**(3), eabc8448 (2021).

130. U. Teğin, P. Wang, and L. V. Wang, "Real-time observation of optical rogue waves in spatiotemporally mode-locked fiber lasers," *Commun. Phys.* **6**(1), 60 (2023).

131. Y. Yang et al., "A diagnostic system toward high-resolution measurement of wavefront profile," *Opt. Commun.* **456**, 124554 (2020).

132. J. C. Jing, X. Wei, and L. V. Wang, "Spatio-temporal-spectral imaging of non-repeatable dissipative soliton dynamics," *Nat. Commun.* **11**(1), 2059 (2020).

133. X. Wei et al., "Real-time frequency-encoded spatiotemporal focusing through scattering media using a programmable 2D ultrafine optical frequency comb," *Sci. Adv.* **6**(8), eaay1192 (2020).

134. J. Park et al., "Snapshot multidimensional photography through active optical mapping," *Nat. Commun.* **11**(1), 5602 (2020).

135. X. Liu et al., "Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging," *Nat. Commun.* **12**(1), 6401 (2021).

136. J. Liang et al., "Homogeneous one-dimensional optical lattice generation using a digital micromirror device-based high-precision beam shaper," *J. Micro/Nanolithogr. MEMS MOEMS* **11**(2), 023002 (2012).

137. Texas-Instruments, "DMD optical efficiency for visible wavelengths (Rev. B)," (2023). https://www.ti.com/lit/an/dlpa083b/dlpa083b.pdf?ts=1701702604909&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FDLP230NP (accessed 23 December 2023).

138. W. Chi and N. George, "Phase-coded aperture for optical imaging," *Opt. Commun.* **282**(11), 2110–2117 (2009).

139. W. Chi and N. George, "Optical imaging with phase-coded aperture," *Opt. Express* **19**(5), 4294–4300 (2011).

140. R. Zhu, T.-H. Tsai, and D. J. Brady, "Coded aperture snapshot spectral imager based on liquid crystal spatial light modulator," in *OSA Tech. Digest, Front. Opt.*, I. Kang and D. Hagan, Eds., p. FW1D.4 (2013).

141. J. Liang, "Introduction to coded optical imaging," in *Coded Optical Imaging*, J. Liang, Ed., Springer Nature (2024).

142. Holoeye, "LC 2012 spatial light modulator (transmissive)," (2023). https://holoeye.com/products/spatial-light-modulators/lc-2012-spatial-light-modulator-transmissive/ (accessed 23 December 2023).

143. J. García-Márquez et al., "Flicker minimization in an LCoS spatial light modulator," *Opt. Express* **20**(8), 8431–8441 (2012).

144. FineLine-Imaging, (2023). https://www.fineline-imaging.com/ (accessed 23 December 2023).

145. J. del Barrio and C. Sánchez-Somolinos, "Light to shape the future: from photolithography to 4D printing," *Adv. Opt. Mater.* **7**(16), 1900598 (2019).

146. M. Marquez et al., "Metalens-based compressed ultra-compact femtophotography: analytical modeling and simulations," *Ultrafast Sci.* **4**, 0052 (2023).

147. J. Yao et al., "Multichannel-coupled compressed ultrafast photography," *J. Opt.* **22**(8), 085701 (2020).

148. A. Matin and X. Wang, "Video encryption/compression using compressive coded rotating mirror camera," *Sci. Rep.* **11**(1), 23191 (2021).

149. A. Matin and X. Wang, "Compressive coded rotating mirror camera for high-speed imaging," *Photonics* **8**(2), 34 (2021).

150. J. Park and L. Gao, "Continuously streaming compressed high-speed photography using time delay integration," *Optica* **8**(12), 1620–1623 (2021).

151. D. Qi et al., "100-trillion-frame-per-second single-shot compressed ultrafast photography via molecular alignment," *Phys. Rev. Appl.* **15**(2), 024051 (2021).

152. G. D. Bai et al., "Multitasking shared aperture enabled with multiband digital coding metasurface," *Adv. Opt. Mater.* **6**(21), 1800657 (2018).

153. Y. Wu et al., "TiO_2 metasurfaces: from visible planar photonics to photochemistry," *Sci. Adv.* **5**(11), eaax0939 (2019).

154. M. Khorasaninejad and F. Capasso, "Metalenses: versatile multifunctional photonic components," *Science* **358**(6367), eaam8100 (2017).

155. X. Hua et al., "Ultra-compact snapshot spectral light-field imaging," *Nat. Commun.* **13**(1), 2732 (2022).

156. T. R. M. Sales and G. M. Morris, "Diffractive–refractive behavior of kinoform lenses," *Appl. Opt.* **36**(1), 253–257 (1997).

157. R. Menon et al., "Maskless lithography," *Mater. Today* **8**(2), 26–33 (2005).

158. P. Wang and R. Menon, "Computational multispectral video imaging," *J. Opt. Soc. Am. A* **35**(1), 189–199 (2018).

159. G. R. B. E. Römer and P. Bechtold, "Electro-optic and acousto-optic laser beam scanners," *Phys. Proc.* **56**, 29–39 (2014).

160. B.-L. Qian and H. E. Elsayed-Ali, "Electron pulse broadening due to space charge effects in a photoelectron gun for electron diffraction and streak camera systems," *J. Appl. Phys.* **91**(1), 462–468 (2002).

161. Y. Wang et al., "Dark current and sensitivity measurements for structured S20 photocathodes," *J. Phys. D: Appl. Phys.* **39**(20), 4341 (2006).

162. M. Liu et al., "Short-wave infrared photoluminescence lifetime mapping of rare-earth doped nanoparticles using all-optical streak imaging," *Adv. Sci.* **11**, 2305284 (2023).

163. V. Degtyareva et al., "Femtosecond streak tubes designing, manufacturing, and testing," in *25th Int. Congr. on High-Speed Photogr. and Photonics* (2003).

164. J. Sasián, *Introduction to Aberrations in Optical Imaging Systems*, Cambridge University Press (2013).

165. Thorlabs, "Small beam diameter scanning galvo mirror systems," (2023). https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=3770&pn=GVS001#4419 (accessed 23 December 2023).

166. H. Wong, Y. L. Yao, and E. S. Schlig, "TDI charge-coupled devices: design and applications," *IBM J. Res. Dev.* **36**(1), 83–106 (1992).

167. W. Uhring et al., "A scalable architecture for multi millions frames per second CMOS sensor with digital storage," in *16th IEEE Int. New Circuits and Syst. Conf. (NEWCAS)*, pp. 252–255 (2018).

168. W. Uhring, J.-B. Schell, and L. Hébrard, "Integrated streak camera with on chip averaging for signal to noise ratio improvement," in *IARIA SIGNAL 2019* (2019).

169. Y. Chiu et al., "Electro-optic beam scanner in KTiOPO$_4$," *Appl. Phys. Lett.* **69**(21), 3134–3136 (1996).

170. K. T. Kim et al., "Photonic streaking of attosecond pulse trains," *Nat. Photonics* **7**(8), 651–656 (2013).

171. J. Chun, H. Jung, and C.-M. Kyung, "Suppressing rolling-shutter distortion of CMOS image sensors by motion vector detection," *IEEE Trans. Consum. Electron.* **54**(4), 1479–1487 (2008).

172. MathWorks, "Control point registration," (2023). https://www.mathworks.com/help/images/control-point-registration.html (accessed 23 December 2023).

173. Y. Lai et al., "Compressed ultrafast tomographic imaging by passive spatiotemporal projections," *Opt. Lett.* **46**(7), 1788–1791 (2021).

174. T. Zhu et al., "Note: x-ray streak camera sweep speed calibration," *Rev. Sci. Instrum.* **81**(5), 056108 (2010).

175. D. Purves, *Brains as Engines of Association: an Operating Principle for Nervous Systems*, Oxford University Press (2019).

176. M. B. Jackson, "Passive current flow and morphology in the terminal arborizations of the posterior pituitary," *J. Neurophysiol.* **69**(3), 692–702 (1993).

177. Y. Gong et al., "Imaging neural spiking in brain tissue using FRET-opsin protein voltage sensors," *Nat. Commun.* **5**(1), 3674 (2014).

178. J. F. Gillooly et al., "Effects of size and temperature on metabolic rate," *Science* **293**(5538), 2248–2251 (2001).

179. X. Zhu et al., "Temperature-feedback upconversion nanocomposite for accurate photothermal therapy at facile temperature," *Nat. Commun.* **7**(1), 10437 (2016).

180. B. B. Lahiri et al., "Medical applications of infrared thermography: a review," *Infrared Phys. Technol.* **55**(4), 221–235 (2012).

181. R. K. Benninger et al., "Quantitative 3D mapping of fluidic temperatures within microchannel networks using fluorescence lifetime imaging," *Anal. Chem.* **78**(7), 2272–2278 (2006).

182. G. Schlegel et al., "Fluorescence decay time of single semiconductor nanocrystals," *Phys. Rev. Lett.* **88**(13), 137401 (2002).

183. S. Kalytchuk et al., "Temperature-dependent exciton and trap-related photoluminescence of CdTe quantum dots embedded in a NaCl matrix: implication in thermometry," *Small* **12**(4), 466–476 (2016).

184. T. Qin et al., "Organic fluorescent thermometers: Highlights from 2013 to 2017," *Trends Anal. Chem.* **102**, 259–271 (2018).

185. T. Chihara et al., "Biological deep temperature imaging with fluorescence lifetime of rare-earth-doped ceramics particles in the second NIR biological window," *Sci. Rep.* **9**(1), 12806 (2019).

186. D. Jaque and F. Vetrone, "Luminescence nanothermometry," *Nanoscale* **4**(15), 4301–4326 (2012).

187. H. Kurokawa et al., "High resolution imaging of intracellular oxygen concentration by phosphorescence lifetime," *Sci. Rep.* **5**(1), 10657 (2015).

188. J. Yang, H. Yang, and L. Lin, "Quantum dot nano thermometers reveal heterogeneous local thermogenesis in living cells," *ACS Nano* **5**(6), 5067–5071 (2011).

189. S. Kiyonaka et al., "Genetically encoded fluorescent thermosensors visualize subcellular thermoregulation in living cells," *Nat. Methods* **10**(12), 1232–1238 (2013).

190. M. Runowski et al., "Upconverting lanthanide doped fluoride NaLuF$_4$: Yb^{3+}-Er^{3+}-Ho^{3+}-optical sensor for multi-range fluorescence intensity ratio (FIR) thermometry in visible and NIR regions," *J. Lumin.* **201**, 104–109 (2018).

191. A. Nexha et al., "Lanthanide doped luminescence nanothermometers in the biological windows: strategies and applications," *Nanoscale* **13**(17), 7913–7987 (2021).

192. O. Otto et al., "Real-time deformability cytometry: on-the-fly cell mechanical phenotyping," *Nat. Methods* **12**(3), 199–202 (2015).

193. J. Jeong et al., "Accurately tracking single-cell movement trajectories in microfluidic cell sorting devices," *PLoS One* **13**(2), e0192463 (2018).

194. D. L. van der Ven et al., "Microfluidic jet impact: spreading, splashing, soft substrate deformation and injection," *J. Colloid Interface* **636**, 549–558 (2023).

195. D. Green, A. Gelb, and G. P. Luke, "Sparsity-based recovery of three-dimensional photoacoustic images from compressed single-shot optical detection," *J. Imaging* **7**(10), 201 (2021).

196. D. Green, A. Gelb, and G. P. Luke, "Compressed single-shot photoacoustic image reconstruction of a 3D pressure distribution," in *OSA Tech. Digest, OSA Imaging and Appl. Opt. Congr. 2021 (3D, COSI, DH, ISA, pcAOP)*, p. CM2E.5 (2021).

197. E. Zhang, J. Laufer, and P. Beard, "Backward-mode multiwavelength photoacoustic scanner using a planar Fabry-Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues," *Appl. Opt.* **47**(4), 561–577 (2008).

198. P. Wang and L. V. Wang, "Single-shot reconfigurable femtosecond imaging of ultrafast optical dynamics," *Adv. Sci.* **10**(13), 2207222 (2023).

199. R. Safaei et al., "High-energy multidimensional solitary states in hollow-core fibres," *Nat. Photonics* **14**(12), 733–739 (2020).

200. Z. Zhu et al., "Attosecond pulse retrieval from noisy streaking traces with conditional variational generative network," *Sci. Rep.* **10**(1), 5782 (2020).

201. W. Luo et al., "Pixel super-resolution using wavelength scanning," *Light Sci. Appl.* **5**(4), e16060 (2016).

202. X. Di et al., "Quantitatively monitoring in situ mitochondrial thermal dynamics by upconversion nanoparticles," *Nano Lett.* **21**(4), 1651–1658 (2021).

203. M. Orozco, "A theoretical view of protein dynamics," *Chem. Soc. Rev.* **43**(14), 5051–5066 (2014).

204. L. Turnbull et al., "Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms," *Nat. Commun.* **7**(1), 11220 (2016).

205. P. Kner et al., "Super-resolution video microscopy of live cells by structured illumination," *Nat. Methods* **6**(5), 339–342 (2009).

206. M. W. Haakestad et al., "Electrically tunable photonic bandgap guidance in a liquid-crystal-filled photonic crystal fiber," *IEEE Photonics Technol. Lett.* **17**(4), 819–821 (2005).

207. M. Ossiander et al., "Extreme ultraviolet metalens by vacuum guiding," *Science* **380**(6640), 59–63 (2023).

208. Q. Cheng et al., "Broadband achromatic metalens in terahertz regime," *Sci. Bull.* **64**(20), 1525–1531 (2019).

209. C. Zuo et al., "Transport of intensity equation: a tutorial," *Opt. Lasers Eng.* **135**, 106187 (2020).

210. F. Zhang et al., "Phase retrieval by coherent modulation imaging," *Nat. Commun.* **7**(1), 13367 (2016).

211. M. Dai et al., "On-chip mid-infrared photothermoelectric detectors for full-Stokes detection," *Nat. Commun.* **13**(1), 4560 (2022).

212. C. Huang et al., "Ultrafast control of vortex microlasers," *Science* **367**(6481), 1018–1021 (2020).

213. B. Heshmat et al., "Photography optics in the time dimension," *Nat. Photonics* **12**(9), 560–566 (2018).

214. Texas-Instruments, "0.47-inch 1080p HSSI DLP® digital micromirror device (DMD)," (2023). https://www.ti.com/product/DLP471NE (accessed 23 December 2023).

Yingming Lai received his BSc degree in optoelectronics from the Southern University of Science and Technology, China, in 2019, and his MSc degree in energy and materials science from Institut National de la Recherche Scientifique (INRS)–Université du Québec, Canada, in 2021. Currently, he is a PhD candidate in the Laboratory of Applied Computational Imaging at INRS. His main research areas are computational imaging, compressive sensing, and ultrafast optical imaging.

Miguel Marquez received his BSc degree in computer science, his MSc degree in applied mathematics, and his PhD in physics from the Universidad Industrial de Santander, Colombia, in 2015, 2018, and 2022, respectively. He is currently a postdoctoral fellow in the Laboratory of Applied Computational Imaging at the Institut National de la Recherche Scientifique (INRS)–Université du Québec, Canada. His main research interests include optical and computational imaging, compressive sensing, high-dimensional signal processing, and optimization algorithms.

Jinyang Liang is an associate professor at the Institut National de la Recherche Scientifique (INRS)–Université du Québec, Canada. He directs the Laboratory of Applied Computational Imaging. He holds Canada Research Chair in Ultrafast Computational Imaging (Tier II). His research interests include ultrafast imaging, computational optics, optical physics, and biophotonics. He received his PhD in electrical engineering from the University of Texas at Austin in 2012. From 2012 to 2017, he was a postdoctoral trainee at Washington University in St. Louis and California Institute of Technology under the supervision of Dr. Lihong V. Wang.