Ultrafast Science
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

Metalens-Based Compressed Ultracompact Femtophotography: Analytical Modeling and Simulations

Miguel Marquez†, Giacomo Balistreri†, Roberto Morandotti, Luca Razzari*, and Jinyang Liang*

Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, Université du Québec, 1650 Boulevard Lionel-Boulet, Varennes, Québec J3X1P7, Canada.

*Address correspondence to: Jinyang.Liang@inrs.ca (J.L.); Luca.Razzari@inrs.ca (L.R.)
†These authors contributed equally to this work.

Citation: Marquez M, Balistreri G, Morandotti R, Razzari L, Liang J. Metalens-Based Compressed Ultracompact Femtophotography: Analytical Modeling and Simulations. *Ultrafast Sci.* 2024;4:Article 0052. https://doi.org/10.34133/ultrafastscience.0052

Submitted 20 June 2023
Accepted 20 November 2023
Published 9 January 2024

Single-shot 2-dimensional optical imaging of transient phenomena is indispensable for numerous areas of study. Among existing techniques, compressed ultrafast photography (CUP) using a chirped ultrashort pulse as active illumination can acquire nonrepetitive time-evolving events at hundreds of trillions of frames per second. However, the bulky size and conventional configurations limit its reliability and application scopes. Superdispersive metalenses offer a promising solution for an ultracompact design with a stable performance by integrating the functions of a focusing lens and dispersive optical components into a single device. Nevertheless, existing metalens designs, typically optimized for the full visible spectrum with a relatively low spectral resolution, cannot be readily applied to active-illumination CUP. To address these limitations, here, we propose single-shot compressed ultracompact femtophotography (CUF) that synergically combines the fields of nanophotonics, optical imaging, compressed sensing, and deep learning. We develop the theory of CUF's data acquisition composed of temporal–spectral mapping, spatial encoding, temporal shearing, and spatiotemporal integration. We also develop CUF's image reconstruction via deep learning. Moreover, we design and evaluate CUF's crucial components—a static binary transmissive mask, a superdispersive metalens, and a 2-dimensional sensor. Finally, using numerical simulations, CUF's feasibility is verified using 2 synthetic scenes: an ultrafast beam sweeping across a surface and the propagation of a terahertz Cherenkov wave.

Introduction

Single-shot compressed ultrafast imaging is a cutting-edge paradigm that allows capturing ultrafast dynamic phenomena occurring in 2-dimensional (2D) space [1]. By combining the power of compressed sensing (CS) and ultrafast imaging, this approach can physically acquire a large number of frames of high-resolution images in a single shot, followed by computational reconstruction to retrieve the movie of the event. Single-shot compressed ultrafast imaging has rapidly attracted substantial interest in physics, biology, chemistry, materials science, and engineering. It provides unprecedented insights into many phenomena, such as shock wave interaction with live cells [2], ultrashort pulse propagation [3], phosphorescence emission evolution [4], instantaneous light scattering [5], and exciton dynamics in photosynthetic light harvesting [6].

Among existing techniques of single-shot compressed ultrafast imaging, a pioneering technique is compressed ultrafast photography (CUP) [7], which works by synergically combining a scheme of 2D spatial encoding using a single static binary transmissive mask with temporal shearing performed by a streak camera with a fully opened entrance slit. As an innovative computational imaging technique, CUP has set a global benchmark for imaging speed, reaching an unprecedented 10 trillion frames per second (Tfps) in passive imaging [8] and >200 Tfps using active illumination [9]. It has a sequence depth (i.e., the number of frames in each recorded movie) of up to 1,000 frames [3] and is compatible with various modalities for microscopic, mesoscopic, and macroscopic imaging. With such an unprecedented real-time imaging speed, CUP is perfectly suited for the accurate observation of many nonrepeatable or difficult-to-repeat complex phenomena and processes formerly inaccessible to pump–probe-based ultrafast imaging, including laser ablation for materials processing [10], light propagation in a scattering medium [11], and optical chaotic dynamics [12]. Leveraging the high universality of its sensing model, the concept of CUP has been implemented in diverse hardware configurations [13]. Meanwhile, deep learning has been increasingly implemented in single-shot compressed ultrafast imaging systems to enhance the image quality and speed in reconstruction, which assists the development of high-level tasks involving complex data and decision-making, the optimization of spatial encoding masks, and the sensing of temporal shearing [14,15].

Within the existing configurations, of our particular interest are CUP systems using an ultrashort chirped pulse as active illumination [9,16–19]. When such a pulse is used to probe a dynamic event, the temporal information is linearly mapped to the spectrum. On the detection side, after spatial encoding, a spatial dispersion element, such as a prism or a grating, shears individual wavelengths to different spatial positions. This modality transfers the burden of temporal shearing to the spectral bandwidth and temporal duration of the illumination pulse. Leveraging the advances in ultrafast laser science, this modality is poised to overcome limitations imposed by state-of-the-art ultrafast optoelectronic devices. Recent developments have pushed the imaging speed to 256 Tfps [8,9,16]. Despite their dazzling performance and high potential, these systems are often limited by their bulky configurations. In particular, temporal shearing and imaging are accomplished using different optical components, which are prone to misalignment and nonlinear shearing from various disruptions (e.g., vibration). Moreover, provided the fixed dispersion ability of gratings or prisms, long-focal-length lenses are needed to guarantee high spectral resolution and hence high imaging speed. Thus, it is challenging to reduce the system size, which limits the use of active-illumination CUP in research and commercial applications, including single-shot pulse characterization in ultrafast laser apparatus [20] and monitoring of ultrafast laser processing of materials [21].

To circumvent these drawbacks, ultracompact snapshot imaging systems (USISs) based on metasurfaces have emerged as a potent approach for high-dimensional (e.g., spatiospectral [22], volumetric [23], multispectral-chiral [24], and spectropolarimetric [25]) data recording in a single exposure. Metasurfaces allow for the flexible control of light properties (e.g., amplitude, phase, and polarization) at subwavelength scales. This advantage has shed light on a wide range of applications requiring high-quality imaging with a restricted system size [26]. Such metasurface-based USISs are highly suitable for snapshot spectral imaging because they are designed to integrate the functions of an imaging lens and a dispersive optical component into a single device [22]. Metasurfaces have also been used as a compact component for intracavity modulation of ultrashort pulses [27], showing their potential to reduce the system size and complexity.

Although burgeoning in optical sensing, metalens-based USISs still have ample room to develop for ultrafast imaging applications. At the device level, most works focused on embedding multifunctionalities of imaging and sensing (e.g., light field, focusing, dispersion, aberration correction, and polarization) in metalenses rather than exploring their potential in compressed coded-aperture imaging techniques [26]. At the system level, existing configurations in metalens-based spectral imaging are typically designed to cover the entire visible spectral range with a spectral resolution of 5 to 10 nm [28,29]. Considering that the bandwidth of a typical femtosecond pulse is 10 to 40 nm, the limited spectral resolution in existing systems makes them impractical for active-illumination CUP.

To address such limitations, we propose single-shot compressed ultracompact femtophotography (CUF)—a new superdispersive metalens-based computational ultrafast imaging modality. Materials and Methods presents CUF's data acquisition model based on the CS theory and its image reconstruction framework based on deep learning. Results and Discussion first describes the design details of 3 crucial components—an encoding mask, a superdispersive metalens, and a 2D sensor—for CUF. Then, it discusses the theoretical validation of CUF via numerical simulations using 2 representative ultrafast scenes. Conclusion summarizes this work and provides a perspective of CUF. Overall, this work provides the first theoretical investigation that verifies the feasibility of metalens-based systems for single-shot real-time ultrafast imaging. The established framework will guide the future prototype construction and experimental validation.

Materials and Methods

Sensing model of CUF

CUF is schematically shown in Fig. 1. Its data acquisition starts by using a linearly chirped ultrashort pulse to probe a dynamic scene denoted by $\mathbf{F} \in \mathbb{R}^{N_x \times N_y \times N_L}$, where N_x and N_y represent the data lengths in the 2 spatial dimensions, and N_L represents the data length in the spectral (hence temporal) dimension. This time-spectrum mapping procedure (denoted by $\mathbf{M} \in \mathbb{R}^{n \times n}$ with $n = N_x N_y N_L$) stores temporal information across the spectral band of the probe pulse [30]. Imaged by the front optics onto a single static (x, y) binary transmissive mask, the transmitted light is spatially encoded (denoted by $\mathbf{C} \in \mathbb{R}^{n \times n}$). Then, the light is relayed by a metalens to an off-axis 2D sensor with unity magnification. The metalens also shears spectral information along a spatial axis (denoted by $\mathbf{S} \in \mathbb{R}^{mN_L \times n}$ with $m = N_x[N_y + (N_L - 1)]$). Finally, the resulting encoded and sheared scene is integrated by the sensor (denoted by $\mathbf{T} \in \mathbb{R}^{m \times mN_L}$), producing a compressively recorded 2D snapshot (denoted by $\mathbf{G} \in \mathbb{R}^{N_x \times [N_y + (N_L - 1)]}$). Overall, CUF's forward model is expressed as

$$\mathbf{g} = \mathbf{\Phi f} + \boldsymbol{\epsilon}. \tag{1}$$

Here, $\mathbf{f} \in \mathbb{R}^{n \times 1}$ is the discrete vector representation of the probe pulse modulated by the dynamic scene $\mathbf{F}$. $\mathbf{g} \in \mathbb{R}^{m \times 1}$ is the vectorial version of the snapshot $\mathbf{G}$. $\boldsymbol{\epsilon} \in \mathbb{R}^{m \times 1}$ represents the noise added to the snapshot during data acquisition. $\mathbf{\Phi} = \mathbf{TSCM} \in \mathbb{R}^{m \times n}$ is the CUF's sensing matrix. More details of CUF's forward model are explained in Note S1.

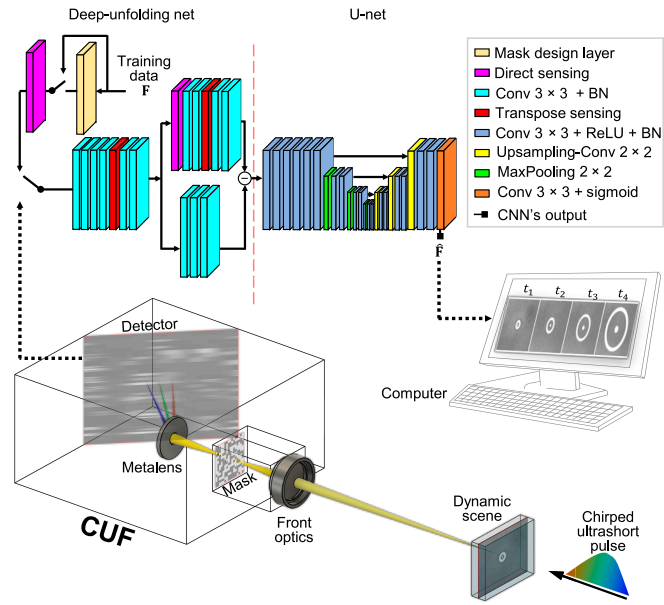

Fig. 1. Design of single-shot compressed CUF using a superdispersive metalens; BN, batch normalization; Conv, convolutional; ReLU, rectified linear unit.

Image reconstruction based on a deep-learning approach

To recover the dynamic scene, the acquired snapshot is input to a convolutional neural network (CNN)-based algorithm composed of 2 cascaded neural networks: a deep-unfolding net that embodies CUF's sensing model and a U-net [31] that takes the output of the first network to improve image reconstruction (Fig. 1). This CNN-based algorithm leverages the merits of both the alternating direction method of multipliers (ADMM) [32] and the network-based CS methods by mapping one iteration of the ADMM steps to a deep network architecture. Manifesting the "splitting-and-optimization" approach embedded in the ADMM, this design is memory efficient and thus essential for learning to reconstruct the (x, y, t) datacube of the dynamic scene (denoted by $\widehat{\mathbf{F}} \in \mathbb{R}^{N_x \times N_y \times N_L}$). In this regard, CUF's image reconstruction is formulated as

$$minimize_{\mathbf{f},\mathbf{z}} \|\mathbf{g} - \mathbf{\Phi f}\|_2^2 + \psi(\mathbf{z}), \quad \text{subject to } \mathbf{z} = \mathbf{f}. \tag{2}$$

Here, $\psi(\cdot): \mathbb{R}^{n\times 1} \rightarrow \overline{\mathbb{R}}$ is a regularizer. $\mathbf{z} \in \mathbb{R}^{n\times 1}$ is an auxiliary variable. To learn the CNN's weights, the loss function $\mathcal{H}(\cdot): \mathbb{R}^{N_x \times N_y \times N_L} \rightarrow \mathbb{R}$ is established as

$$\mathcal{H}\left(\mathbf{F}, \widehat{\mathbf{F}}\right) = l_1\left(\mathbf{F}, \widehat{\mathbf{F}}\right) + l_{\text{SSIM}}\left(\mathbf{F}, \widehat{\mathbf{F}}\right) + l_1\left(\mathbf{G}, \mathcal{G}\left(\widehat{\mathbf{F}}\right)\right), \tag{3}$$

where $\mathcal{G}$ is the direct sensing operator, $l_1(\cdot)$ is the l_1-norm, and $l_{\text{SSIM}}(\cdot)$ represents the structural similarity (SSIM) index. A detailed explanation of the image reconstruction algorithm is included in Note S2.

Results and Discussion

Design and evaluation of key components in CUF

Mask

The mask, as the first of the 3 essential components of CUF, is responsible for spatial encoding. In CUF, a static binary transmissive mask is used to passively encode the dynamic scene, which allows a compact design. We designed the mask by using a customized layer that incorporated a sigmoid function $R_{i_x,i_y} = (1 + e^{-\theta_{i_x,i_y}})^{-1}$, where $\theta_{ix,iy}$ ($i_x = \{0, \ldots, N_x - 1\}$ and $i_y = \{0, \ldots, N_y - 1\}$) represent the layer's weights and were initialized with random values. Interfacing CUF's CNN (see Fig. 1), this layer molded the mask structure, which was guided by the CUF's sensing geometry and the spatiotemporal features of the training data. A constraint function [33] (denoted by $\mathcal{Q}(\cdot): \mathbb{R}^{N_x \times N_y} \rightarrow \mathbb{R}$), whose minima were obtained only when $R_{ix,iy}$ were 0 or 1, was used to generate the binary mask

$$\mathcal{Q}(\mathbf{R}) = \sum_{i_x=0}^{N_x-1} \sum_{i_y=0}^{N_y-1} \left(R_{i_x,i_y}\right)^2 \left(R_{i_x,i_y} - 1\right)^2. \tag{4}$$

As shown in Fig. 2A, the designed mask has several advantages compared to the conventional pseudo-random binary mask (detailed in Note S3). First, it reduced the content in the low-frequency range (Fig. 2B) [34]. As a result, the cross-correlation of any 2 different columns obtained an average correlation coefficient of 0.47 (Fig. 2C), showing better linear independence against the random mask with an average correlation coefficient of 0.49. Moreover, its structure enabled the acquisition of compressed measurements with a more uniform pixel intensity distribution and a higher dynamic range (Fig. S1).

The mask's encoding pixel size, determined by the metalens diameter and focal length (see Metalens) as well as the sensor pixel size (see Sensor), was fixed at 11.6 μm. Finally, the size of 512 × 512 encoding pixels was used to limit the CUF's field of view (FOV), ensuring a full capture of the compressively recorded 2D snapshot by the sensor.

Metalens

We designed a polarization-insensitive superdispersive metalens made of Si_3N_4 (refractive index n_{Si3N4}) meta-atoms on a SiO_2 (refractive index n_{SiO2}) substrate to accommodate the specifications of a frequency-doubled ytterbium femtosecond laser (central wavelength of $\lambda_c = 515$ nm; spectral bandwidth of $\Delta\lambda = 16$ nm [35]). The hyperbolic phase profile of the superdispersive metalens exhibited a wavelength-dependent focal length (f) [36]. Among possible meta-atom shapes (e.g., parallelepipeds [37–39], cylinders [40,41], and elliptic cylinders [42]), we used polarization-insensitive cylinders arranged in a square lattice (Fig. 3A), which exploited variations in the meta-atoms radius (r_c) to obtain a phase coverage of 2π [29,30].

To determine the optimal metalens configuration, we conducted a parameter sweep by varying meta-atom period (Λ), r_c, and thickness (d) (marked in Fig. 3A) with 3 restrictions. First, Λ must be smaller than the light wavelength in SiO_2 (i.e.,

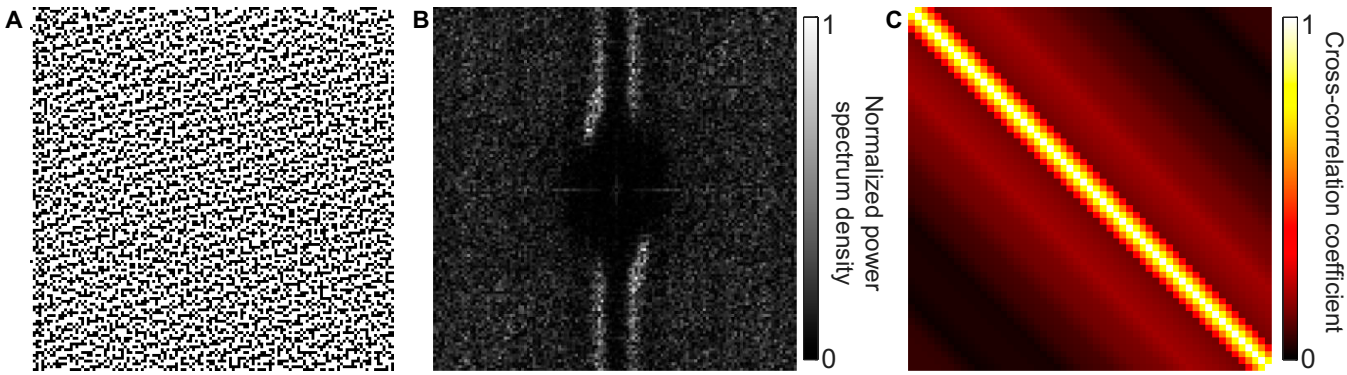

Fig. 2. Design of the encoding mask for CUF. (A) Designed binary mask. (B) Fourier spectrum of (A). The dc component was set to zero for display purposes. (C) Two-dimensional cross-correlation map of columns in (A).

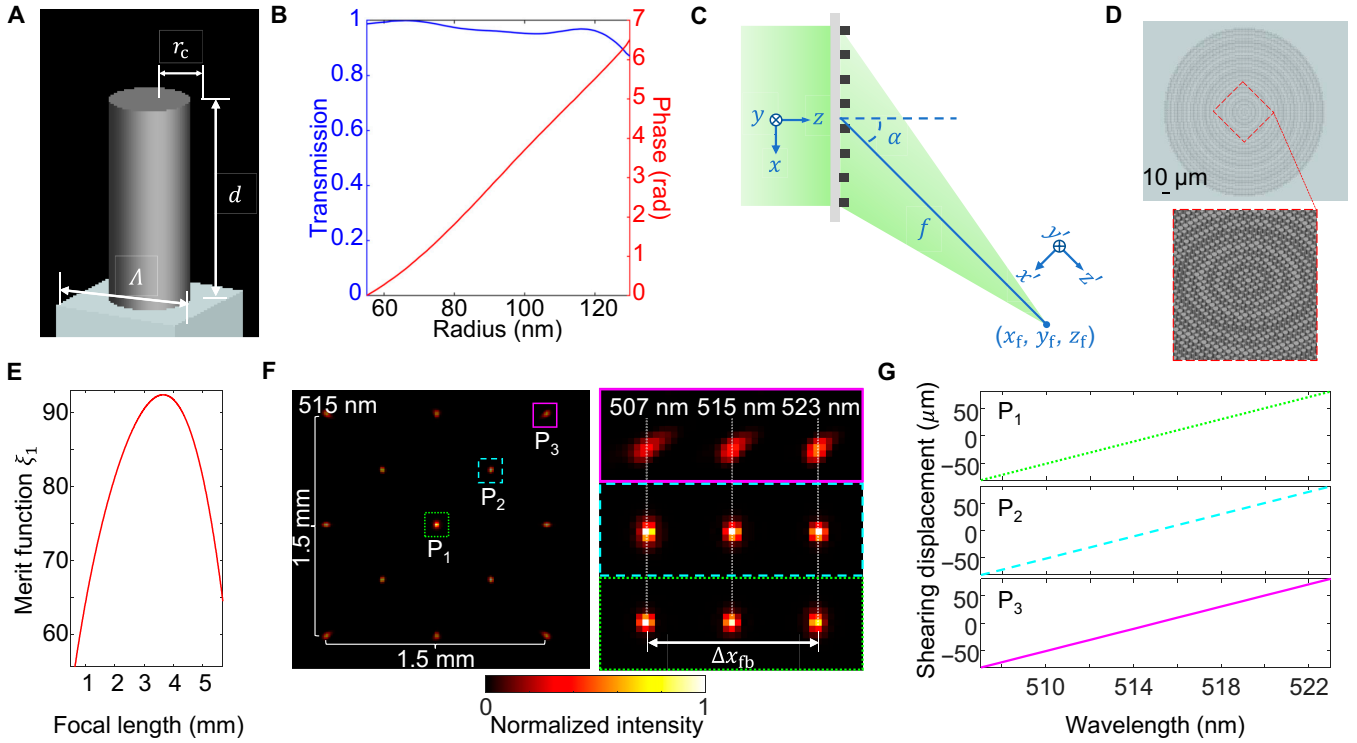

Fig. 3. Design of the superdispersive metalens for CUF. (A) Schematic of the cylinder meta-atom employed for the design of the metalens, where r_c, d, and Λ are the meta-atom radius, thickness, and meta-atom period, respectively. (B) Amplitude transmission (blue line) and phase profile (red line) of the cylinder meta-atom with r_c varying from 50 to 130 nm. (C) Working principle of the metalens, where f and α are the focal length and the focusing angle, respectively. (x_f, y_f, z_f) are the coordinates of the focus. (D) Top view (top panel) of the designed metalens and 3D zoomed-in view of its central region (bottom panel). (E) Merit function, ξ_1, of the simulated metalenses as a function of the focal length. (F) Left panel: Representative PSFs produced by the designed metalens over a FOV of 1.5 mm × 1.5 mm. Right panels: Zoomed-in views of the PSFs at 3 wavelengths from selected locations (marked by the colored boxes in the left panel). (G) Relation between the wavelengths and the relative shearing displacement of 3 selected locations marked in (F).

$\Lambda < \lambda_c/n_{SiO2}$). Second, d must be chosen so that the phase modulation covers the full 2π range [41]. Third, the minimum value of the cylinder aspect ratio (i.e., $2r_c/d$) was limited to 1/8, for compatibility with commonly used metasurface fabrication techniques (such as electron-beam lithography [41]). The behavior of the meta-atoms was simulated using a finite-difference time-domain algorithm (Ansys Lumerical) to solve Maxwell's equations for the propagating electromagnetic field. This analysis led to the optimized values of $\Lambda = 270$ nm, $d = 760$ nm, and $r_c = 50$ to 130 nm. The results show an almost constant transmittance of ~90% for all the values of r_c, and 2π coverage of the phase profile for the entire targeted spectral band (i.e., $\lambda_c =$ 515 nm and $\Delta\lambda = 16$ nm). As an example, the corresponding transmission and phase profile for the central wavelength are illustrated in Fig. 3B.

Using these parameters, we designed a superdispersive lens producing spectral shearing along the x' direction in the focal plane (Fig. 3C and D). The lens' phase profile was constructed by using the meta-atoms with a suitable r_c according to the radius-phase mapping shown in Fig. 3B. We simulated a lens that focused light at an angle $\alpha = \sin^{-1}(x_f/f)$, where x_f is the x' coordinate of the focus. The diameter of the simulated metalens was set to 200 μm, a limit imposed solely by the computational capacity of our workstation. The dispersion angle was chosen to be $\alpha =$ 45° for $\lambda_c = 515$ nm, so that light beams at different wavelengths could be detected and well resolved on the camera plane.

As the next step, we simulated several metalenses exhibiting different focal lengths ($f = 0.5$ to 5.5 mm) to obtain a good focusing efficiency (FE) and have a large shearing distance in the x' direction for the considered spectral range ($\Delta x_{fb} = x_{f@523} - x_{f@507}$). Here, FE is defined as the ratio of the total intensity in a circular aperture at the focal plane with a radius 3 times the full width at half maximum (FWHM) of the focal spot to the total intensity of the light transmitted through the lens. Because FE and Δx_{fb} showed opposite trends with the increase of the focal length, we looked for the focal length value that maximizes a merit function defined by their product (i.e., $\xi_1 = \text{FE} \bullet \Delta x_{fb}$). As shown in Fig. 3E, the evaluation pinpoints the optimized focal length of $f =$ 3.705 mm, at the wavelength λ_c.

Finally, using the designed metalens, we retrieved the point spread functions (PSFs) on the focal plane for the targeted spectral range in a FOV of 1.5 mm × 1.5 mm (Fig. 3F). PSFs from representative positions in the FOV are shown in the left panel of Fig. 3F. At the center, the FWHMs of the PSFs and the shearing distance were evaluated to be 10.5 ± 0.5 μm (mean ± standard deviation) and $\Delta x_{fb} = 162.7$ μm, resulting in a dispersion coefficient of $\mu = 10.15\ \mu\text{m} \cdot \text{nm}^{-1}$. From the center to the periphery, the PSF's FWHM increased to 16.3 ± 1.5 μm, which was attributed to the increased propagation distance. Notably, the shearing remained the same, as shown in Fig. 3G. These results showed low spatial distortion in PSFs and nearly spatial invariance in the shearing distance. All these characteristics were built into CUF's sensing model to accurately reconstruct the ultrafast video (see details in Notes S1 and S2).

Sensor

The 2D sensor is the last essential component of CUF, responsible for capturing a compressed snapshot of the dynamic scene.

Sensor specifications can affect the compressed image quality, frame rate, compression ratio, and hence both spatial and temporal resolution, as well as the sequence depth. The frame rate and the sequence depth of the CUF system are derived from

$$r = \frac{\mu}{|\eta| d_s}, \tag{5}$$

and

$$N_L = \left\lceil \frac{\Delta\lambda \cdot \mu}{d_s} \right\rceil, \tag{6}$$

respectively. Here, η is the chirping parameter of the probe pulse, d_s is the sensor pixel size along the temporal shearing direction, and $(\Delta\lambda \cdot \mu)$ is the sheared distance. Furthermore, the FOV in CUF is determined by the size of the encoding mask, the sensor, the sequence depth, and the overall magnification ratio of the system.

To investigate the effectiveness of CUF, we carried out an ablation analysis of how the sensor pixel size and spectral beam separation affect the reconstruction performance (Fig. 4). For this analysis, we used the specifications of 3 off-the-shelf sensors with pixel sizes of $d_s = 2.9$ μm (Sensor 1, Starlight IMX462-99 IR-CUT with 1,920 × 1,080 pixels), $d_s = 3.4$ μm (Sensor 2, FLIR BFS-GE-16S2C-BD2 with 1,440 × 1,080 pixels), and $d_s = 5.8$ μm (Sensor 3, FLIR GS3-U3-23S6M-C with 1,920 × 1,200 pixels). Then, to examine the spectral shearing operation, we selected 8 bandwidths of $\Delta\lambda = \{4, 6, 8, 9, 10, 12, 14, 16\}$ nm. The FOV for the 3 sensor configurations was limited by the mask to a value of 1.5 mm × 1.5 mm. Finally, to promote a fair comparison in the training and reconstruction, we defined a baseline database generated from Sensor 1 specifications, which was used to construct the databases for Sensor 2 and Sensor 3. The baseline database was created from the "Challenge on Spectral Reconstruction from RGB Images" database [43]. To train the CNN weights, we randomly selected and cropped 800 datacubes with $N_x \times N_y = 512 \times 512$ pixels (for Sensor 1), $N_x \times N_y = 430 \times 430$ pixels (for Sensor 2), and $N_x \times N_y = 256 \times 256$ pixels (for Sensor 3). N_L for all 3 sensors (with the 8 selected bandwidths) were determined via Eq. 6. All simulations were implemented in Tensorflow [44] and trained on Intel Core i9 12900K 3.2GHz with a GeForce RTX 3090 GPU (32GB RAM) using the ADAM optimizer [45].

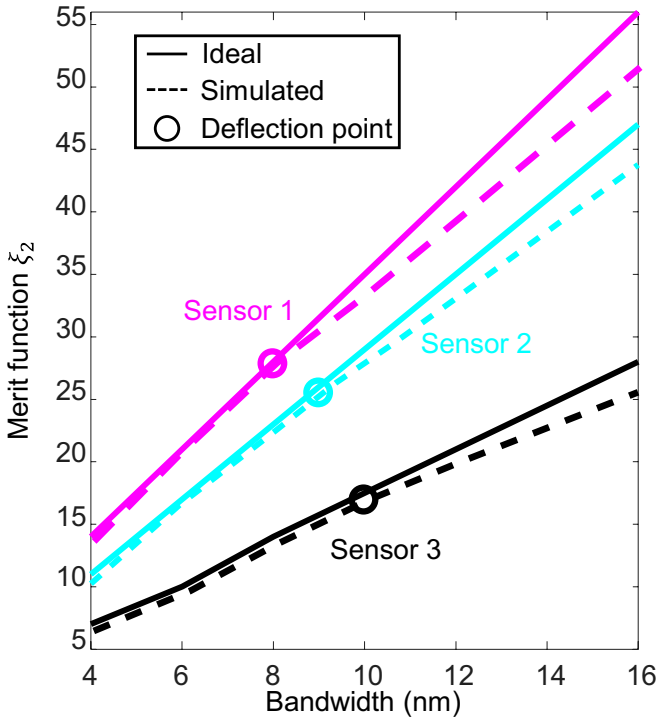

Fig. 4. Merit function, ξ_2, across 3 sensors over different illumination bandwidths for the density of 21.6 lp/mm.

We established a merit function as the product of the modulation transfer function (MTF) and the sequence depth (i.e., $\xi_2 = \text{MTF} \times N_L$). For this analysis, we designed a negative line-pair target with a density of 21.6 lp/mm as a static target. The reconstructed result was compared to the ideal case (i.e., MTF = 1). As shown in Fig. 4, this comparison reveals high reconstruction accuracy (i.e., $\xi_2 \geq 0.98 N_L$) for the scenarios with $N_L \leq 27$ frames (i.e. $\Delta\lambda = 8$ nm) for Sensor 1, $N_L \leq 25$ frames (i.e., $\Delta\lambda = 9$ nm) for Sensor 2, and $N_L \leq 17$ frames (i.e., $\Delta\lambda = 10$ nm) for Sensor 3. Beyond this deflection point, the merit metric decreases by up to 9%. This result showed that Sensor 1, which had the smallest pixel size, could recover sharper spatial features. However, the fixed CNN's ability to handle spatial and temporal information indicated a trade-off between the (x, y) image size (i.e., $N_x \times N_y$) and the sequence depth (i.e., N_L), which was demonstrated by the tolerance of Sensor 3 to higher bandwidths with a lower merit function. Additional details of image reconstruction ablation are provided in Note S4.

Feasibility demonstration via simulation

CUF of a beam sweeping across a surface

To demonstrate CUF's feasibility in ultrafast imaging, we conducted simulation validations using 2 representative scenes. First, we emulated CUF in a dynamic scene where a temporally chirped pulse with a front tilt shone on a negative resolution target with line pairs whose densities ranged from 19.2 to 172.4 lp/mm. The incident pulse could be generated by combining temporal (e.g., using a glass rod) and spatial (e.g., using a grating pair [16]) chirps. For this simulation, we created a synthetic dataset with a FOV of 1.5 mm × 1.5 mm and a pixel size $d_s = 2.9$ μm (Sensor 1). Using the chirp parameter in [20] (i.e., $\eta = 65.6$ fs • nm^{-1}), a bandwidth of $\Delta\lambda = 16$ nm, and a temporal Gaussian profile, we simulated an imaging speed of $r = 53.3$ Tfps and a sequence depth of $N_L = 56$ frames. The resulting training database comprises 900 videos with $N_x \times N_y \times N_L = 512 \times 512 \times 56$ pixels. Based on Eq. 1, the spatially encoded and temporally sheared scene was spatiotemporally integrated to form a compressed measurement with a size of up to 512 × 567 pixels. After data acquisition, the deep-learning reconstruction algorithm recovers a datacube with the loss function defined in Eq. 3.

The proposed CUF methodology reconstructed this transient scene. Five exemplary frames of the scene (as the ground truth) and their corresponding frames reconstructed by CUF are shown in Fig. 5A. The full reconstruction (Movie S1) shows swift sweeping of the illumination across this target. To further analyze CUF's spatial reconstruction accuracy, the ground truth and the reconstruction were temporally integrated (Fig. 5B). The comparison of 6 representative local regions with different line-pair densities shows good accuracy in CUF's reconstruction for the line pairs with a density of up to 49 lp/mm. Moreover, to quantify CUF's applicability, we tested 100 scenes

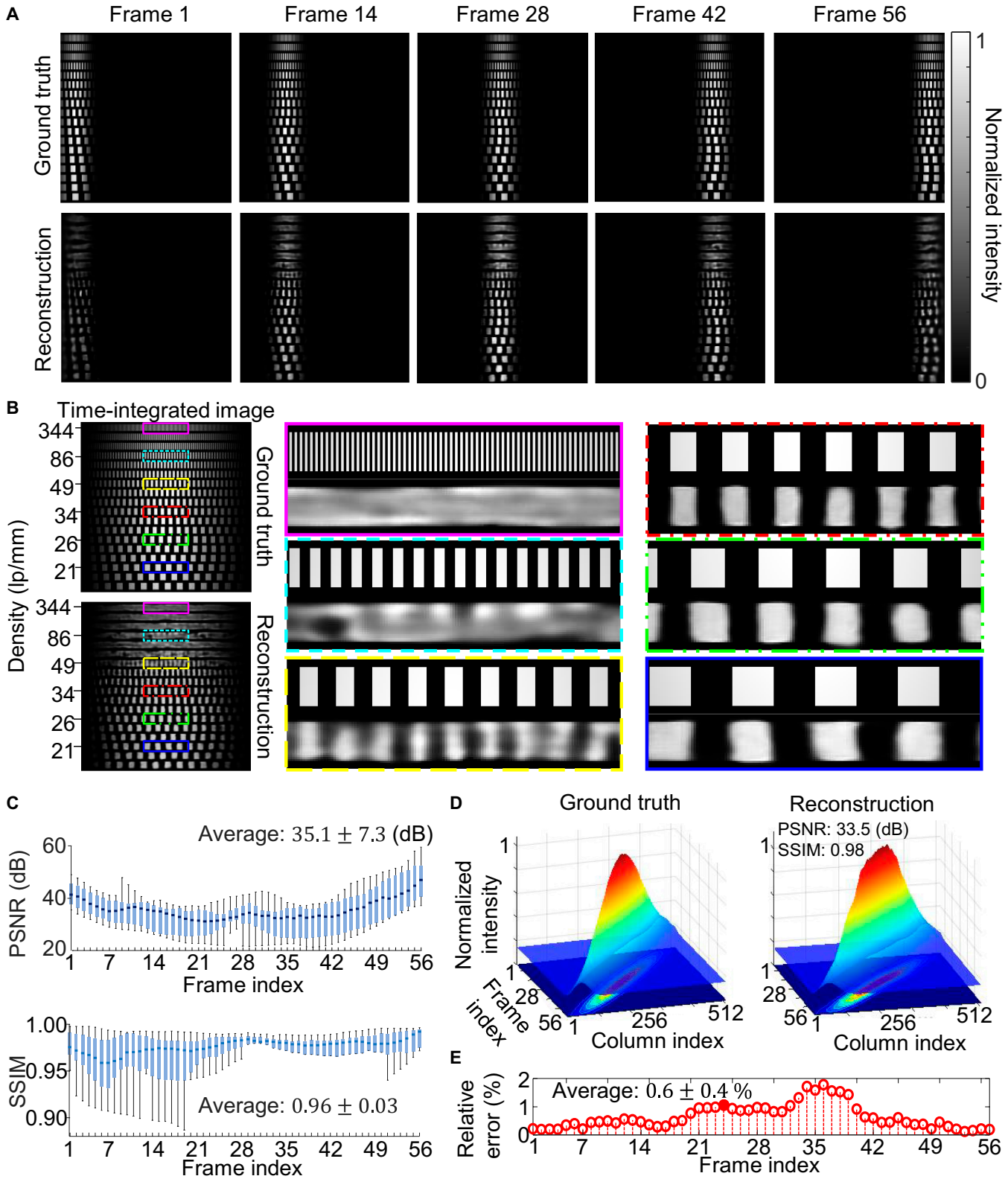

Fig. 5. Simulation of CUF for a spatiotemporally chirped ultrashort laser pulse transmitted through a line-pair target. (A) Five selected frames of the ground truth and the CUF reconstruction. (B) Time-integrated images and comparison of 6 representative line-pair targets with different densities. (C) Statistical analysis of peak PSNRs and SSIM indices of the reconstructed results using 100 testing datacubes. (D) Comparison of the estimated illumination profile with the ground truth. (E) Time course of the relative error of reconstructed horizontal position of illumination center.

that had the same illumination condition in Fig. 5A but different patterns. Shown in Fig. 5C, the results reflect a high averaged peak signal-to-noise ratio (PSNR) of 35.1 ± 7.3 dB and SSIM index of 0.96 ± 0.03. Finally, to demonstrate the CUF's ability to retrieve information on the sample illumination, we averaged each frame in the vertical direction for both the

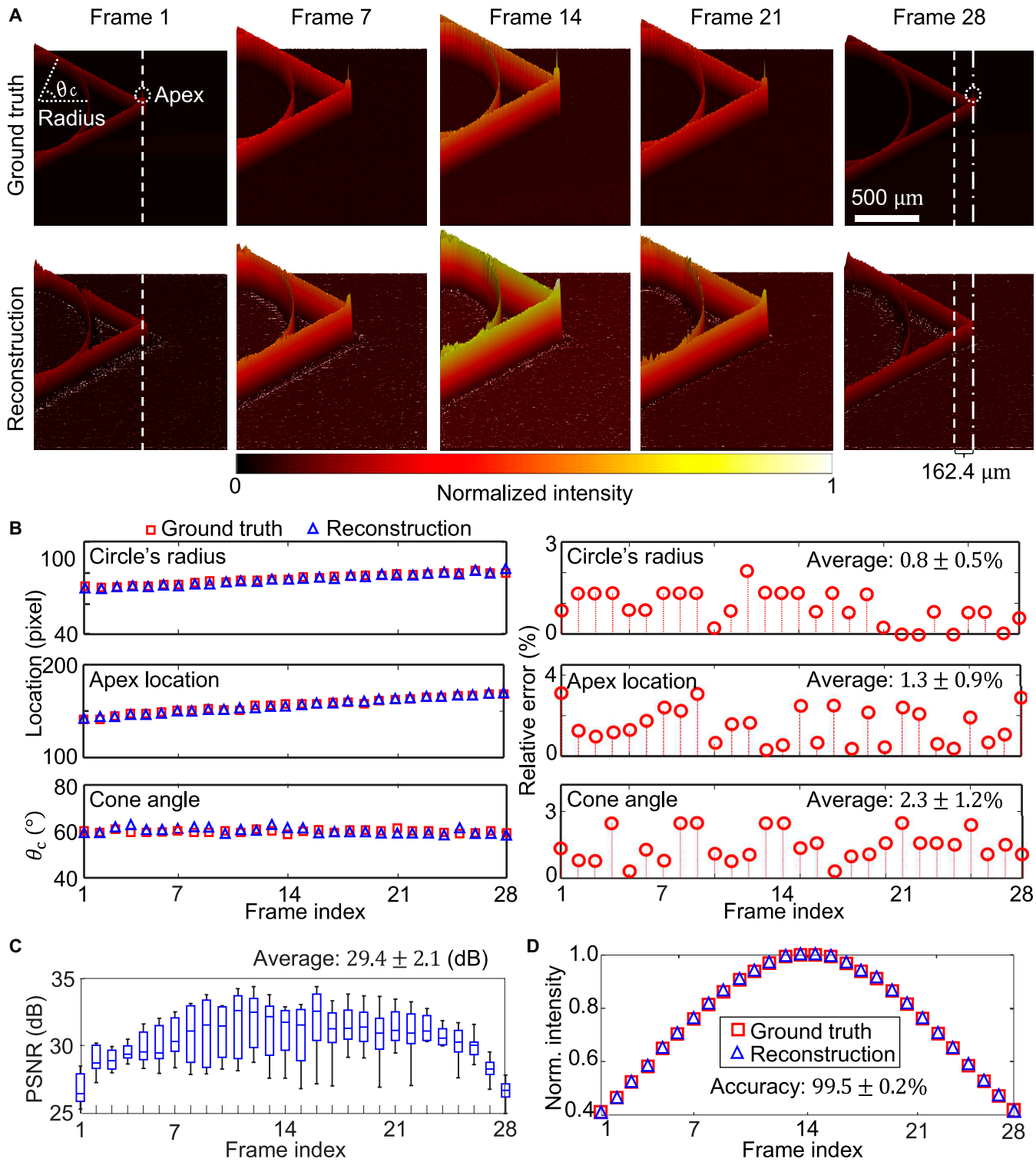

Fig. 6. Simulation of CUF of a propagating Cherenkov wave. (A) Five selected frames of the ground truth and reconstruction. The apex locations in Frame 1 and Frame 28 are marked by the white dashed line and the white dashed-dotted line, respectively. (B) Left column: Time courses of the reconstructed circle's radius, apex location, and cone angle between the ground truth and the CUF reconstruction. Right column: Time courses of relative errors of the results shown in the left column. (C) Statistical analysis of PSNRs from 100 emulated test scenes. (D) Comparison of the estimated spectral intensity profile from the CUF reconstruction in (A) with the ground truth.

ground truth and the reconstruction. The line profile of the reconstruction was then divided by that of the ground truth for each nonzero-intensity pixel. In this way, the pulse front tilt in the illumination is shown in Fig. 5D. The reconstruction shows excellent quality with 33.5 dB in PSNR and 0.98 in SSIM. Using this result, the time course of relative errors of the horizontal position of the illumination center was calculated, which yielded an average error of 0.6 ± 0.4% (Fig. 5E).

CUF of a propagating terahertz Cherenkov wave

In the second ultrafast phenomenon simulation, we emulated CUF of a propagating Cherenkov wave in a nonlinear crystal. This phenomenon could be generated by focusing a near-infrared, millijoule-level, femtosecond laser pulse into a bulk lithium niobate ($LiNbO_3$) single crystal [46]. The generated terahertz Cherenkov wave could be imaged using a temporally linearly chirped probe beam propagating perpendicularly to the Cherenkov wave propagation direction. We created a synthetic dataset by using the experimental conditions mentioned in [46]. In particular, the focused laser pulse had a propagation speed of 1.37×10^8 m·s^{-1}, and the terahertz Cherenkov wave had a cone angle of $\theta_c = 60.0 \pm 1.0°$. We simulated CUF of this event at an imaging speed of $r = 23.6$ Tfps using Sensor 3 (i.e., $d_s = 5.8$ μm) and the chirp parameter $\eta = 74.1$ fs·nm^{-1}. The resulting training database comprises 1,800 videos with $N_x \times N_y \times N_L = 256 \times 256 \times 28$ pixels. To promote an accurate reconstruction of the high-spatial-frequency components in the structure of the Cherenkov wave, the loss function was slightly modified to $\mathcal{H}\left(\mathbf{F},\hat{\mathbf{F}}\right) = l_1\left(\mathcal{R}(\mathbf{F}),\mathcal{R}\left(\hat{\mathbf{F}}\right)\right) + l_{SSIM}\left(\mathcal{R}(\mathbf{F}),\mathcal{R}\left(\hat{\mathbf{F}}\right)\right) + l_1\left(\mathbf{G},\mathcal{G}\left(\hat{\mathbf{F}}\right)\right)$, where $\mathcal{R}(\cdot) = \left|\mathcal{F}_{2D}^{-1}\{\mathcal{F}_{2D}\{\cdot\} \otimes \mathbf{B}\}\right|$ with $\mathcal{R}(\cdot): \mathbb{R}^{N_x \times N_y \times N_L} \to \mathbb{R}^{N_x \times N_y \times N_L}$ Here, $\mathcal{F}_{2D}\{\cdot\}: \mathbb{C}^{N_x \times N_y \times N_L} \to \mathbb{C}^{N_x \times N_y \times N_L}$ and $\mathcal{F}_{2D}^{-1}\{\cdot\}: \mathbb{C}^{N_x \times N_y \times N_L} \to \mathbb{C}^{N_x \times N_y \times N_L}$ are the direct and inverse 2D discrete Fourier transform operators, respectively; $\mathbf{B} \in \mathbb{R}^{N_x \times N_y \times N_L}$ represents the binary circular high-pass filter with a diameter of 125 μm; $\otimes$ represents the element-wise Hadamard product.

CUF showed its feasibility to accurately reconstruct the datacube. As shown in Fig. 6A, 6 selected frames of the ground truth are compared with their corresponding reconstructed frames. The full movie is shown in Movie S2. To quantitatively analyze the reconstruction, Fig. 6B presents the time evolution of the radius of the transition-like radiation (associated with the spherical wave first observed in [46]), the centroid on the cone apex, and the cone angle. The average error in the Cherenkov cone geometry estimation was calculated to be 2.3%. Moreover, we generated 100 emulated Cherenkov radiation videos, where the starting spatial location of the cone apex and the cone angle were randomly varied. Figure 6C shows the time course of the reconstructions with an averaged PSNR of 29.7 ± 2.3 dB. To quantitatively analyze CUF's ability to measure the illumination features, Fig. 6D presents the precise matching between the emulated spectral intensity profile to the one estimated directly from the CUF's reconstruction. The accuracy estimation in the spectral intensity profile was calculated to be 99.5 ± 0.2%. These results show CUF's robustness in accurately sensing ultrafast phenomena and recovering the illumination spectral profile with high accuracy.

Conclusion

Our analytical modeling and simulation have shown that the synergy of a chirped ultrashort pulse, a superdispersive metalens, CS, and neural network-based image reconstruction can realize intelligent ultracompact ultrafast imaging. CUF compressively records a dynamic scene in a 2D snapshot, which is then processed by a CNN that uses the framework of ADMM to promote a higher resonance between the PSF-based CUF's sensing model and the training of the CNN's weights. The set of 2 parallel layers is built to model the direct sensing and the transpose process operations involved in CUF's inverse problem. The design of 3 crucial components in CUF is described in detail. A transmissive binary mask is designed to promote a uniform pixel intensity distribution and a higher dynamic range. For the metalens, polarization-insensitive cylinder meta-atoms are used to produce full 2π phase coverage with ~90% transmittance for all the investigated radii values. Along with the encoding mask, the metalens-engineered dispersive response allows the mixture of temporal and spatial information in the compressively recorded snapshot. Finally, CUF is successfully evaluated using 3 different sensor configurations with pixel sizes ranging from 2.9 to 5.8 μm. Numerical simulations demonstrate high-quality reconstruction using the CUF paradigm in 2 representative ultrafast dynamic scenes, which provides the first feasibility verification of metalens-aided compressed ultrafast imaging.

This study lays the theoretical foundation for the future development of the CUF prototype. In particular, the chirped ultrashort laser pulse of the targeted spectral bandwidth can be generated by a frequency-doubled ytterbium femtosecond laser [35]. The metalens can be realized by first depositing a Si_3N_4 film using plasma-enhanced chemical vapor deposition and then patterning it via electron-beam lithography [47]. The designed mask can be customized using high-resolution printing [4]. Finally, the sensor can be chosen from commercially available off-the-shelf products. Along with the short focal length of the designed metalens (i.e., $f = 3.705$ mm), the availability of existing techniques provides a viable path to the development of an ultracompact prototype. It is envisaged that CUF will provide a versatile platform with a scalable spatial resolution by coupling various imaging modalities from microscopes to telescopes and with various charge-coupled device/complementary metal-oxide semiconductor cameras. Leveraging on the simultaneous use of shearing and imaging provided by the metalens and the superior reconstructed image quality offered by the deep-learning algorithms, CUF is expected to be implemented and improve the reliability and stability of various femtophotography applications, such as in imaging longitudinal ultrafast all-optical switching [48], sensing irreversible chemical reaction dynamics of organic crystals [49], and evolution of plasma wakes in a laser wakefield accelerator [50].

Acknowledgments

Funding: This work was supported in part by Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-05959, RGPAS-2017-507845, I2IPJ-555593-20, RGPIN-2018-06217, RGPAS-2018-522650, and RGPIN-2019-06138); Canada Foundation for Innovation and Ministère de l'Économie et de l'Innovation du Québec (37146); Fonds de Recherche du Québec–Nature et Technologies (203345–Centre d'Optique, Photonique, et Lasers); and Canada Research Chairs Program (CRC-2022-00119).

Author contributions: J.L. proposed the conceptual CUF system. M.M. and G.B. developed the CUF's image reconstruction algorithm and performed the simulations. G.B., R.M., L.R., and J.L. designed the metalens. M.M. and J.L. analyzed the data. M.M., G.B., and J.L. drafted the manuscript. All authors revised the manuscript. J.L. and L.R. initiated and supervised the project.

Competing interests: The authors declare that they have no competing interests.

Data Availability

The data reported in this article is available upon request to the authors.

Supplementary Materials

Figs. S1 to S2
Notes S1 to S4
Movies S1 to S2

References

1. Liang J. Punching holes in light: Recent progress in single-shot coded-aperture optical imaging. *Rep Prog Phys*. 2020;83(11):Article 116101.
2. Wood JC, Chapman DJ, Poder K, Lopes NC, Rutherford ME, White TG, Albert F, Behm KT, Booth N, Bryant JSJ, et al. Ultrafast imaging of laser driven shock waves using betatron x-rays from a laser wakefield accelerator. *Sci Rep*. 2018;8:11010.
3. Liang J, Zhu L, Wang LV. Single-shot real-time femtosecond imaging of temporal focusing. *Light: Sci Appli*. 2018;7(1):42.
4. Liu X, Skripka A, Lai Y, Jiang C, Liu J, Vetrone F, Liang J. Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging. *Nat Commun*. 2021;12(1):6401.
5. Liang J, Ma C, Zhu L, Chen Y, Gao L, Wang LV. Single-shot real-time video recording of a photonic Mach cone induced by a scattered light pulse. *Sci Adv*. 2017;3:Article e1601814.
6. Malý P, Gruber JM, Cogdell RJ, Mančal T, Van Grondelle R. Ultrafast energy relaxation in single light-harvesting complexes. *Proc Natl Acad Sci*. 2016;113(11):2934–2939.
7. Gao L, Liang J, Li C, Wang LV. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature*. 2014;516(7529):74–77.
8. Tang H, Men T, Liu X, Hu Y, Su J, Zuo Y, Li P, Liang J, Downer MC, Li Z. Single-shot compressed optical field topography. *Light: Sci Appli*. 2022;11:244.
9. Wang P, Wang LV. Single-shot reconfigurable femtosecond imaging of ultrafast optical dynamics. *Adv Sci*. 2023;10(13):Article e2207222.
10. Mao X, Mao S, Russo RE. Imaging femtosecond laser-induced electronic excitation in glass. *Appl Phys Lett*. 2003;82(5):697–699.
11. Barber B, Putterman S. Observation of synchronous picosecond sonoluminescence. *Nature*. 1991;352(6333):318–320.
12. Merritt D, Milosavljević M, Verde L, Jimenez R. Dark matter spikes and annihilation radiation from the galactic center. *Phys Rev Lett*. 2001;88(19):Article 191301.
13. Lai Y, Xue Y, Côté CY, Liu X, Laramée A, Jaouen N, Légaré F, Tian L, Liang J. Single-shot ultraviolet compressed ultrafast photography. *Laser Photonics*. 2020;14(10):2000122.
14. Marquez M, Lai Y, Liu X, Jiang C, Zhang S, Arguello H, Liang J. Deep-learning supervised snapshot compressive imaging enabled by an end-to-end adaptive neural network. *IEEE J Sel Top Sig Proc*. 2022;16(4):688–699.
15. Zhang A, Wu J, Suo J, Fang L, Qiao H, Li DDU, Zhang S, Fan J, Qi D, Dai Q, et al. Single-shot compressed ultrafast photography based on U-net network. *Opt Express*. 2020;28(26):39299–39310.
16. Wang P, Liang J, Wang LV. Single-shot ultrafast imaging attaining 70 trillion frames per second. *Nat Commun*. 2020;11:2091.
17. Lu Y, Wong TT, Chen F, Wang L. Compressed ultrafast spectral-temporal photography. *Phys Rev*. 2019;122(19): Article 193904.
18. Touil M, Idlahcen S, Becheker R, Lebrun D, Rozé C, Hideur A, Godin T. Acousto-optically driven lensless single-shot ultrafast optical imaging. *Light: Sci Appli*. 2022;11(1):66.
19. Xie C, Meyer R, Froehly L, Giust R, Courvoisier F. In-situ diagnostic of femtosecond laser probe pulses for high resolution ultrafast imaging. *Light: Sci Appli*. 2021;10(1):126.
20. Steffen B, Gerth C, Caselle M, Felber M, Kozak T, Makowski DR, Mavrič U, Mielczarek A, Peier P, Przygoda K, et al. Compact single-shot electro-optic detection system for THz pulses with femtosecond time resolution at MHz repetition rates. *Rev Sci Instrum*. 2020;91(4):Article 045123.
21. Malinauskas M, Žukauskas A, Hasegawa S, Hayasaki Y, Mizeikis V, Buividas R, Juodkazis S. Ultrafast laser processing of materials: From science to industry. *Light: Sci Appli*. 2016;5(8):e16133–e16133.
22. Hua X, Wang Y, Wang S, Zou X, Zhou Y, Li L, Yan F, Cao X, Xiao S, Tsai DP, et al. Ultra-compact snapshot spectral light-field imaging. *Nat Commun*. 2022;13:2732.
23. Chen C, Song W, Chen JW, Wang JH, Chen YH, Xu B, Chen MK, Li H, Fang B, Chen J, et al. Spectral tomographic imaging with aplanatic metalens. *Light: Sci Appli*. 2019;8(1):99.
24. Khorasaninejad M, Chen WT, Zhu AY, Oh J, Devlin RC, Rousso D, Capasso F. Multispectral chiral imaging with a metalens. *Nano Lett*. 2016;16(7):4595–4600.
25. Lin Z, Pestourie R, Roques-Carmes C, Li Z, Capasso F, Soljačić M, Johnson SG. End-to-end metasurface inverse design for single-shot multi-channel imaging. *Opt Express*. 2022;30(16):28358–28370.
26. Pan M, Fu Y, Zheng M, Chen H, Zang Y, Duan H, Li Q, Qiu M, Hu Y. Dielectric metalens for miniaturized imaging systems: Progress and challenges. *Light: Sci Appli*. 2022;11(1):195.
27. Divitt S, Zhu W, Zhang C, Lezec HJ, Agrawal A. Ultrafast optical pulse shaping using dielectric metasurfaces. *Science*. 2019;364(6443):890–894.
28. Holsteen AL, Lin D, Kauvar I, Wetzstein G, Brongersma ML. A light-field metasurface for high-resolution single-particle tracking. *Nano Lett*. 2019;19(4):2267–2271.
29. Yang Z, Wang Z, Wang Y, Feng X, Zhao M, Wan Z, Zhu L, Liu J, Huang Y, Xia J, et al. Generalized Hartmann-shack array of dielectric metalens sub-arrays for polarimetric beam profiling. *Nat Commun*. 2018;9:4607.
30. Diels JC, Rudolph W. *Ultrashort laser pulse phenomena*. Albuquerque (NM): Elsevier; 2006.
31. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Paper presented at: Medical Image Computing and Computer-Assisted Intervention (MICCAI); 2015 Oct 5–9; Munich, Germany.
32. Yuan X, Liu Y, Suo J, Dai Q. Plug-and-play algorithms for large-scale snapshot compressive imaging, Paper presented at: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020 Jun 13–19. Seattle, WA.
33. Bacca J, Galvis L, Arguello H. Coupled deep learning coded aperture design for compressive image classification. *Opt Express*. 2020;28(6):8528–8540.
34. Lau DL, Arce GR. *Modern digital halftoning*. Boca Raton (FL): CRC Press; 2018.
35. Rovere A, Jeong YG, Piccoli R, Lee SH, Lee SC, Kwon OP, Jazbinsek M, Morandotti R, Razzari L. Generation of high-field terahertz pulses in an HMQ-TMS organic crystal pumped by an ytterbium laser at 1030 nm. *Opt Express*. 2018;26(3):2509–2516.
36. Khorasaninejad M, Chen WT, Oh J, Capasso F. Super-dispersive off-axis meta-lenses for compact high resolution spectroscopy. *Nano Lett*. 2016;16(6):3732–3737.
37. Khorasaninejad M, Chen WT, Devlin RC, Oh J, Zhu AY, Capasso F. Metalenses at visible wavelengths: Diffraction-

limited focusing and subwavelength resolution imaging. *Science*. 2016;352(6290):1190–1194.
38. Groever B, Rubin NA, Mueller JB, Devlin RC, Capasso F. High-efficiency chiral meta-lens. *Sci Rep*. 2018;8:7240.
39. Liang H, Lin Q, Xie X, Sun Q, Wang Y, Zhou L, Liu L, Yu X, Zhou J, Krauss TF, et al. Ultrahigh numerical aperture metalens at visible wavelengths. *Nano Lett*. 2018;18(7):4460–4466.
40. Khorasaninejad M, Zhu AY, Roques-Carmes C, Chen WT, Oh J, Mishra I, Devlin RC, Capasso F. Polarization-insensitive metalenses at visible wavelengths. *Nano Lett*. 2016;16(11):7229–7234.
41. Fan ZB, Shao ZK, Xie MY, Pang XN, Ruan WS, Zhao FL, Chen YJ, Yu SY, Dong JW. Silicon nitride metalenses for close-to-one numerical aperture and wide-angle visible imaging. *Phys Rev Appli*. 2018;10(1):Article 014005.
42. Arbabi E, Arbabi A, Kamali SM, Horie Y, Faraon A. High efficiency double-wavelength dielectric metasurface lenses with dichroic birefringent meta-atoms. *Opt Express*. 2016;24(16):18468–18477.
43. Arad B, Timofte R, Ben-Shahar O, Lin YT, Finlayson GD. Ntire 2020 challenge on spectral reconstruction from an rgb image. Paper presented at: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020 Jun 13–19. Seattle, WA.
44. Zaccone G, Karim MR. *Deep Learning with TensorFlow: Explore neural networks and build intelligent systems with Python*. Birmingham (UK): Packt Publishing Ltd; 2018.
45. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imag*. 2016;3(1):47–57.
46. Wang Z, Su F, Hegmann FA. Ultrafast imaging of terahertz Cherenkov waves and transition-like radiation in LiNbO3. *Opt Express*. 2015;23(6):8073–8086.
47. Wang C, Yu Z, Zhang Q, Sun Y, Tao C, Wu F, Zheng Z. Metalens eyepiece for 3D holographic near-eye display. *Nano*. 2021;11(8):1920.
48. Kuttruff J, Garoli D, Allerbeck J, Krahne R, de Luca A, Brida D, Caligiuri V, Maccaferri N. Ultrafast all-optical switching enabled by epsilon-near-zero-tailored absorption in metal-insulator nanocavities. *Commun Phys*. 2020;3(1):114.
49. Poulin PR, Nelson KA. Irreversible organic crystalline chemistry monitored in real time. *Science*. 2006;313(5794):1756–1760.
50. Sävert A, Mangles SPD, Schnell M, Siminos E, Cole JM, Leier M, Reuter M, Schwab MB, Möller M, Poder K, et al. Direct observation of the injection dynamics of a laser wakefield accelerator using few-femtosecond shadowgraphy. *Phys Rev Lett*. 2015;115(5):Article 055002.