Tailoring Space-Time Nonlocality for Event-Based Image Processing Metasurfaces

Sedigheh Esfahani,[1,2,*] Michele Cotrufo,[1,3,*] and Andrea Alù[1,2,†]
[1]Photonics Initiative, Advanced Science Research Center, City University of New York, New York, New York 10031, USA
[2]Physics Program, Graduate Center of the City University of New York, New York, New York 10016, USA
[3]The Institute of Optics, University of Rochester, Rochester, New York 14627, USA

(Received 9 October 2023; accepted 27 June 2024; published 7 August 2024)

Analog computation with passive optical components can enhance processing speeds and reduce power consumption, recently attracting renewed interest thanks to the opportunities enabled by metasurfaces. Basic image processing tasks, such as spatial differentiation, have been recently demonstrated based on engineered nonlocalities in metasurfaces, but next-generation computational schemes require more advanced capabilities. Here, by simultaneously tailoring the nonlocal electromagnetic response of a metasurface in space and time, we demonstrate a passive ultrathin silicon-based device that performs mixed spatiotemporal differentiation of input images, realizing event-based edge detection. The metasurface performs spatial differentiation only when the input image is evolving in time, resulting in spatiotemporal image processing on subpicosecond timescales. Moreover, the metasurface design can be tailored to selectively enhance objects moving at desired speeds. Our results point towards fully passive processing of spatiotemporal signals, for highly compact neuromorphic cameras.

DOI: 10.1103/PhysRevLett.133.063801

Designing faster and energy-efficient tools for data processing is of paramount importance due to the exponentially growing rate of data creation, processing and storage [1]. Replacing digital processing with analog optical computing has recently attracted renewed interest [2–5] due to the appealing possibility of manipulating data at the speed of light while avoiding analog-to-digital conversion [6]. Image processing is among the most important computational tasks, with critical applications in computer vision [7], neuromorphic computing [8,9], artificial intelligence [10], and augmented reality [11].

In the last decade, analog passive computing has evolved from bulky elements to ultrathin structured layers—known as metasurfaces [12]—which are compatible with large-area fabrication and prone to miniaturization and integration [13]. Several works [14–22] have demonstrated that optical metasurfaces can perform different image processing tasks, such as edge detection [14–19,23], *without requiring a 4F system* [24,25], with the potential of drastically reducing the footprint of all-optical analog computers. In this approach, a metasurface with tailored angle-dependent transmission acts as a spatial-momentum filter, thus performing a desired spatial operation—e.g., differentiation—by filtering the spatial Fourier components of the input image. In this context, *nonlocal* metasurfaces [14–16,26–28] (whose optical response is dictated by the coherent interaction between many unit cells) supporting engineered Fano resonances and/or quasibound states in the continuum have been instrumental to tailor nontrivial angle-dependent responses, leading to efficient image-processing devices.

While most studies on metamaterial-based analog computation have so far focused on *spatial* operations, comparatively little effort has been devoted to the implementation of *temporal* and, more broadly, *spatiotemporal* computation. Similarly to 4F systems for spatial operations, temporal operations can be performed via bulky pulse shapers, whereby the frequency components of a pulse are separated by diffraction gratings and filtered by masks [29,30]. Notably, the working principles of metamaterial-based spatial image processing can be extended to the temporal domain: by tailoring its *frequency* response, a metasurface can act as a compact spectral filter, thus performing time-domain computations *without* using pulse shapers and diffraction gratings. Following this strategy, previous theoretical works have investigated temporal processing with graphene [31] or dielectric [32] structures, including recent experimental demonstration of time differentiation at near-visible wavelengths [33] and radio frequencies [34]. Recent studies have also investigated metasurfaces that perform linear combinations of spatial and temporal differentiations [35–38], to generate spatiotemporal optical vortices [39,40].

Besides signal processing, performing analog and high-speed spatiotemporal computations can largely benefit [41] the emerging field of neuromorphic computing (NC) [8,9]—a computational framework where algorithms and hardware mimic the human brain. Rather than capturing information at

*These authors contributed equally to this work.
†Contact author: aalu@gc.cuny.edu

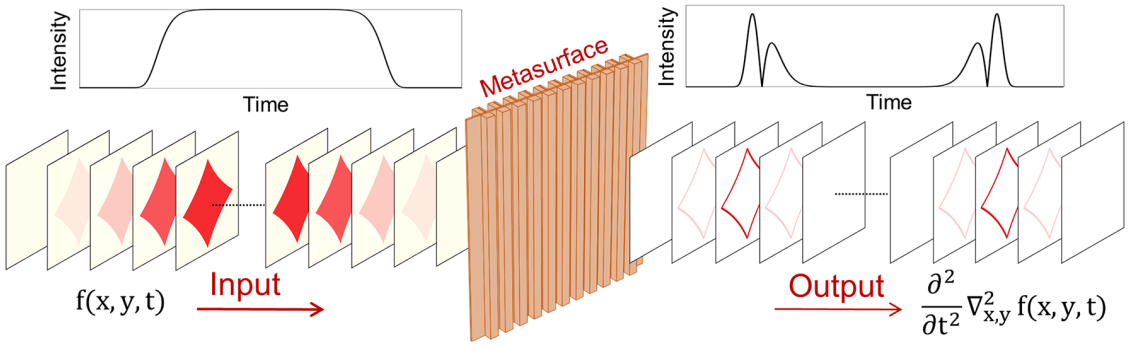

FIG. 1. Schematic of spatiotemporal edge detection with metasurfaces. Left: A spatiotemporal input signal $f(x, y, t)$, encoded in the envelope of a traveling wave, impinges on the metasurface (center). Right: The output signal corresponds to the space-time derivative of the input.

regular intervals, *neuromorphic* or *event-based* cameras respond to local changes of the input signal in time [42]. Current implementations of neuromorphic cameras rely on sophisticated circuitry that detects changes in brightness between neighboring pixels [43,44], triggering data acquisition. These approaches are affected by limited temporal and spatial resolution and large latency, and they require active components and voltage bias, preventing miniaturization and energy savings. Notably, the basic neuromorphic principle—a spatial computation triggered by a temporal change—can be mimicked by *mixed* spatiotemporal differentiation [e.g., $f(x, y, t) \rightarrow \partial_t^2 \nabla_{x,y}^2 f(x, y, t)$], as outlined in this Letter. Yet, implementing these operations with analog metasurfaces is challenging, as it requires engineering a highly nontrivial frequency-momentum transfer function.

In this work, we show that dispersion engineering of nonlocal metasurfaces can be extended to the spatiotemporal domain, simultaneously tailoring angular and frequency responses. In particular, we demonstrate a general recipe to design metasurfaces performing cascaded spatiotemporal differentiation. Based on this principle, we design ultrathin passive metasurfaces, with a relatively simple geometry, performing spatiotemporal *mixed* second-order differentiation [e.g., $\partial_t^2 \nabla_{x,y}^2$] on an input signal. This operation performs analog, all-optical, and fully passive *event-based edge detection*: only regions of the input signal with simultaneous nonzero spatial and time gradients are transmitted. In other words, the spatial edges of an input image are selectively transmitted *only when their local intensity changes in time*, as sketched in Fig. 1: a spatiotemporal signal $f(x, y, t)$, encoded in the envelope of a traveling electromagnetic wave, impinges on a metasurface with tailored spatiotemporal nonlocality. For graphical simplicity, in Fig. 1 we consider a simplified case of a factorized input signal, i.e., $f(x, y, t) = f_t(t) \times f_{x,y}(x, y)$ where $f_{x,y}(x, y)$ is a binary-valued 2D image (the red rhombus), and the overall intensity changes in time following a pulselike function $f_t(t)$ (top-left of Fig. 1). The metasurface filters this signal, creating an output signal proportional to $\partial_t^2 \nabla_{x,y}^2 f(x, y, t)$, whereby the spatial edges of the input are enhanced only at times when the input intensity is simultaneously varying. We emphasize that such functionality can only be implemented with metasurfaces performing mixed spatiotemporal operations [e.g., $\partial_t^2 \nabla_{x,y}^2$], and it cannot be achieved with devices performing linear combinations of temporal and spatial derivatives [35–38]. We numerically demonstrate that this advanced computational task can be achieved in a simple, fully passive, dielectric subwavelength metasurface. Compared to event-based electronic NCs, our approach is not affected by finite temporal resolution and latency, and it involves no energy consumption. Moreover, the subwavelength footprint of the proposed metasurface allows for easy integration in compact devices.

We begin by discussing how a general spatiotemporal differential operator can be implemented with nonlocal metasurfaces (MSs). For simplicity, we focus on 1D images, but the concept can be extended to 2D. A general time-varying image, carried by a structured electromagnetic wave with carrier frequency ω_0, is described by $\tilde{f}(x,t) = e^{i\omega_0 t} f(x,t) = e^{i\omega_0 t} \int d\Omega dk_x e^{-i[k_x x - \Omega t]} F(k_x, \Omega)$, where $f(x, t)$ is a slowly varying (in time) envelope function, $F(k_x, \Omega)$ is its Fourier transform, and $\Omega \equiv \omega - \omega_0$. When the image impinges on a linear MS, the transmitted image is determined by the filtering of spatiotemporal Fourier components, i.e., by the MS momentum-frequency transfer function (TF) $t(k_x, \Omega)$. In particular, a mixed spatiotemporal derivative $(\partial^m/\partial x^m)(\partial^n/\partial t^n) f(x, t)$ is obtained if the TF is $t^{(m,n)}(k_x, \Omega) = k_x^m \Omega^n$ (up to an overall phase factor). Because the second-order differentiation is better suited for isotropic edge detection [16,27], here we focus on the operator $(\partial^2/\partial x^2)(\partial^2/\partial t^2)$, but our approach can be generalized to other operators. Our target response corresponds to the TF $t^{(2,2)}(k_x, \Omega) = k_x^2 \Omega^2$ [Fig. 2(c)], which is achieved when a MS supports a transmission zero at $\Omega = 0$ and normal incidence ($k_x = 0$) and, additionally, the

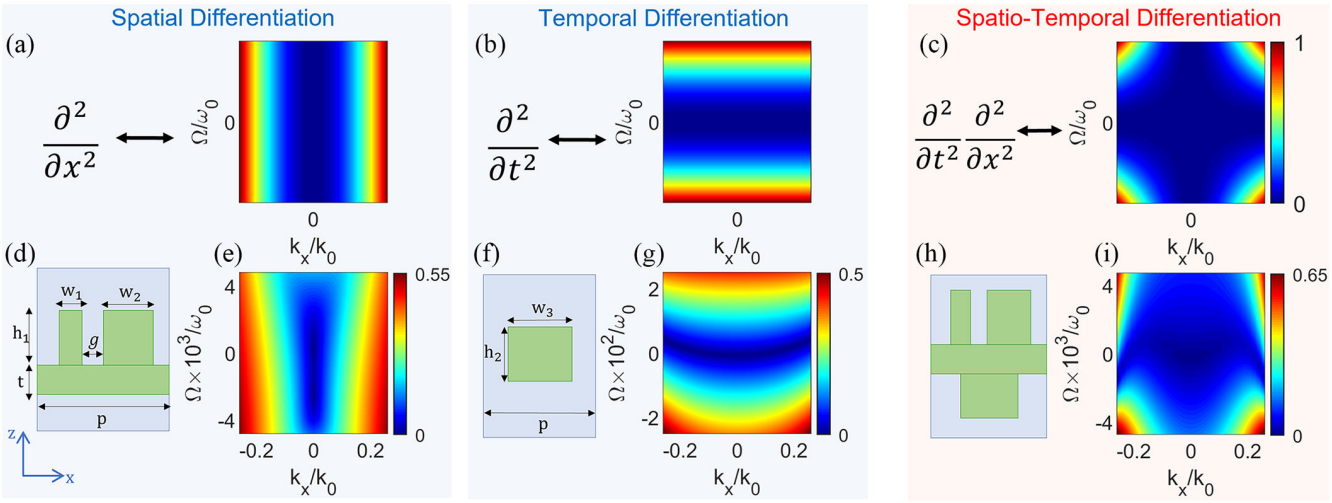

FIG. 2. Ideal transfer function (TF), geometry, and numerical simulations of the proposed MS. (a)–(c) Magnitude of the ideal TF to perform (a) second-order spatial derivative, (b) second-order time derivative, and (c) mixed spatiotemporal derivative. (d)–(i) Unit cells and numerically calculated TFs of realistic metasurfaces performing the corresponding operations shown on top of each box. The parameters are $p = 810$, $h_2 = 393$, $w_3 = 430$, $h_1 = 540$, $w_1 = 80$, $w_2 = 260$, $g = 55$, and $t = 83$ nm. The geometrical dimensions of the metasurface in panel h are provided in [46].

transmission increases only when *both* Ω and k_x are simultaneously varied, while it remains zero when either Ω or k_x is zero. Such nontrivial TF is markedly different from the one necessary to implement linear combinations of spatial and temporal derivatives [35–38]. Designing MSs with arbitrary spatiotemporal nonlocal TFs is generally nontrivial. Inverse-design methods can help this task [4,32], and spatiotemporal coupled-mode theories can provide powerful guidelines [38,45].

Here, we use a more intuitive approach, whereby the required spatiotemporal response is obtained by engineering the spatial and temporal nonlocalities of two cascaded MSs. Indeed, the TF $t^{(2,2)}(k_x,\Omega)$ can be realized as the product of the TF of a MS performing spatial-only differentiation [$(t^{(2,0)}(k_x,\Omega) = k_x^2$, Fig. 2(a)] and the TF of a MS performing temporal-only differentiation [$(t^{(0,2)}(k_x,\Omega) = \Omega^2$, Fig. 2(b)]. Following this intuition, we optimize two different MSs that separately perform these two operations [14,34]. For concreteness, we assume silicon-based 1D MSs embedded in air (Fig. 2), and we target an operational wavelength $\lambda \approx 1500$ nm$[(\omega_0/2\pi) \approx 200$ THz]. All electromagnetic simulations in Fig. 2 were performed with Comsol. The MSs are orthogonal to the z direction and uniform along y, and the impinging electric field is y polarized. Figure 2(d) shows the geometry of a 1D MS whose TF, $t_s^{\mathrm{MS}}(k_x,\Omega)$, closely approximates the TF required for second-order spatial derivative, $t^{(2,0)}(k_x,\Omega)$ (see Fig. 2 caption for the geometry). We note that the design in Fig. 2(d) allows breaking horizontal and/or vertical mirror symmetries. Such broken symmetries are not fundamentally required to implement the even operator $(\partial^2/\partial x^2)$, but they would be required for odd-order derivatives [48,49]. Figure 2(f), instead, shows the unit cell of another MS whose TF, $t_t^{\mathrm{MS}}(k_x,\Omega)$, is close to the one implementing second-order temporal differentiation $[t^{(0,2)}(k_x,\Omega)]$. In first approximation, i.e., neglecting near-field perturbations and feedback due to back-and-forth reflections, cascading these two MSs leads to a device with a TF given by the product $t_s^{\mathrm{MS}}(k_x,\Omega) \times t_t^{\mathrm{MS}}(k_x,\Omega)$, realizing the desired spatio-temporal derivation. In practice, multiple reflections are non-negligible, especially in this scenario where each individual MS is highly reflective at normal incidence. Nonetheless, since the two MS independently target nonlocality in orthogonal spaces, and they are designed to be nondispersive in the dual coordinate, cascading the two designs shown in Figs. 2(d),2(f) provides a useful first-order design to obtain the desired functionality (see also [46]). The residual feedback due to multiple reflections is minimized by adjusting the distance between the MSs to prevent resonances, and by finely tuning the geometrical parameters to account for spectral shifts induced by feedback. Remarkably, for this particular device we found that, even when the two MSs are cascaded back to back without any gap between them, after fine-tuning their geometry, the resulting metasurface [Fig. 2(h)] supports a TF [denoted $t_{st}^{\mathrm{MS}}(k_x,\Omega)$, Fig. 2(i)] that agrees well with the target one [Fig. 2(c)]. However, we emphasize that such zero-gap configuration might not apply to other devices and functionalities. The phases of all TFs shown in Fig. 2 are reported in [46].

While we were able to design a single-layer MS [Fig. 2(h)] with a TF [Fig. 2(i)] close to ideal [Fig. 2(c)], small discrepancies arise. In particular, $t_{st}^{\mathrm{MS}}(k_x,\Omega)$ is not symmetric along the frequency axis, e.g., the high-transmission lobes at high frequencies are less pronounced

than the ones at lower frequencies. In order to assess these discrepancies more quantitatively, we expanded the calculated TF $t_{st}^{\mathrm{MS}}(k_x,\Omega)$ [Fig. 2(i)] over the basis $t^{(m,n)}(k_x,\Omega)$. This analysis, discussed in [46], reveals that the weight of the term $t^{(2,2)}(k_x,\Omega)$ in the expansion of $t_{st}^{\mathrm{MS}}(k_x,\Omega)$ is larger than 93% of the combined total weights, confirming that $t_{st}^{\mathrm{MS}}(k_x,\Omega)$ mainly performs the $(\partial^2/\partial x^2)(\partial^2/\partial t^2)$ operation. Moreover, the desired response is obtained over a broad range of wavevectors and frequencies: The spatial numerical aperture ($\mathrm{NA_s}$), which defines the maximum spatial frequency of the input image that the MS can process, is $\mathrm{NA_s} = k_{x,\mathrm{max}}/k_0 = \sin(\theta_{\mathrm{max}}) \approx 0.26 (\theta_{\mathrm{max}} \approx 15°)$, corresponding to a maximum spatial resolution of $\mathrm{R_s} = (\lambda/2\mathrm{NA_s}) \approx 3$ μm. Similarly, we define a temporal numerical aperture, $\mathrm{NA_t} = \Omega_{\mathrm{max}}/\omega_0 \approx 0.005$, which quantifies the maximum spectral bandwidth of the input signal. The metasurface temporal resolution (i.e., the smallest time separation that can be resolved in the input signal), is $\mathrm{R_t} = \pi/(\omega_0 \mathrm{NA_t}) \approx 0.5$ ps. The ratio of these two numerical apertures defines the *characteristic speed* $\mathrm{v_0} \equiv \mathrm{cNA_t/NA_s}$ (where c is the speed of light), an important parameter of the device discussed more below.

We now show how this MS can effectively implement spatiotemporal edge detection. We first consider the case where an input image composed of spatial segments of different lengths, whose intensities are abruptly switched between zero and one [Fig. 3(a)] is processed by the metasurface in Figs. 2(h) and 2(i). The spatial and temporal dimensions of the input are adjusted to fit the maximum spatial ($\mathrm{NA_s}$) and temporal ($\mathrm{NA_t}$) frequencies of the designed MS, respectively. As expected, in the calculated [46] output image [Fig. 3(b)] only the *spatiotemporal* edges—i.e., the areas where the input image features strong gradients both in space and time, corresponding to the corners of the rectangles in Fig. 3(a)—are largely enhanced with respect to other areas of the signal. In other words, the MS highlights spatial edges only when the edges are *simultaneously* evolving in time. For comparison, Fig. 3(c) shows the output image calculated using the ideal TF [Fig. 2(c)], which exactly implements the $(\partial^2/\partial x^2)(\partial^2/\partial t^2)$ operation. The close resemblance between Figs. 3(b) and 3(c) confirms that the MS mainly performs the targeted spatiotemporal differentiation. Minor detrimental effects are also visible in Fig. 3(b): the spatial-only or temporal-only edges are partially transmitted (although with weaker intensity than the spatio-temporal edges), due to non-idealities of the TF in Fig. 2(i) [46].

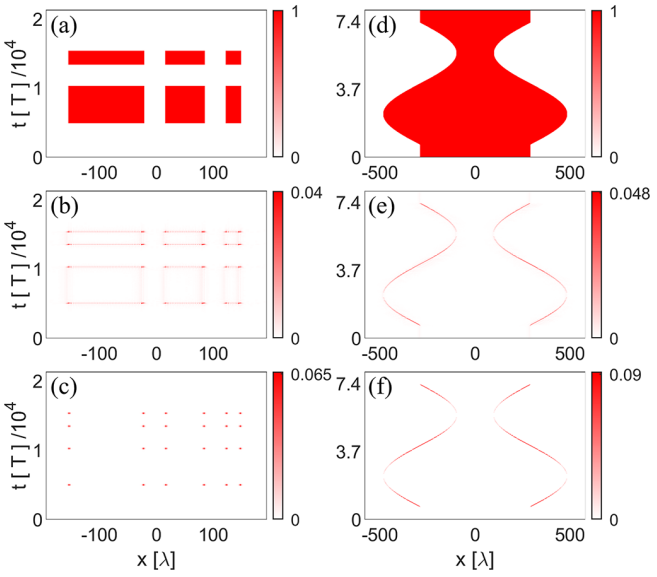

FIG. 3. Spatiotemporal intensity of input and output signals. (a) Time-dependent 1D image consisting of segments with fixed size and whose intensity is switched on and off in time. (b)–(c) Corresponding output image, calculated using (b) the realistic transfer function [Fig. 2(i)] and the ideal transfer function [Fig. 2(c)]. (d)–(f) Same as in panels (a)–(c), but with a different spatiotemporal input, consisting of a single segment whose length changes in time. Times are in units of $T = (2\pi/\omega_0) \approx 4.5$ fs and lengths are in units of $\lambda \approx 1.5$ μm.

The designed MS also features a high throughput efficiency: the peak intensity transmitted in the metasurface-filtered image [Fig. 3(b)] is over 60% of the intensity obtained in the ideal scenario [Fig. 3(c)]. Such efficiencies are comparable to the efficiencies experimentally achieved in spatial-only edge detection [16] and temporal differentiation [33]. Next, we show that our MS can also process input signals that are not factorized in spatial and temporal parts. In Fig. 3(d), we consider a single segment with fixed intensity and whose width changes in time. As expected, in the filtered output image [Fig. 3(e)] the edges of the segment are enhanced only at times when the segment is undergoing a width change—i.e., when the effective position of a spatial edge is changing in time. Once again, the output image filtered by the metasurface [Fig. 3(e)] perfectly matches the image obtained by applying the exact spatiotemporal differentiation [Fig. 3(f)].

The proposed MS can be used for event-based edge detection, that is, it enhances the edges of an object—whose image is projected on the metasurface—only when the object is moving. In Fig. 4(a), we consider an input image made of a segment with fixed width which is initially at rest and then starts moving. In the first part of the motion, the image maintains a constant speed equal to 0.5 v_0, where v_0 is the characteristic speed defined above. The speed then increases in discrete steps, first to v_0 and then to 1.5 v_0, as indicated by the color-coded shaded areas in Fig. 4(a). Then, it undergoes a uniformly accelerated motion [gray shaded area in Fig. 4(a)], until it stops again. In the filtered spatiotemporal signal [Fig. 4(b)], the edges of the image are enhanced only when the image is moving, as expected. Moreover, the edge intensities clearly depend on the image speed, with the largest intensity achieved when the image is moving at a speed close to v_0. This effect can be explained by considering how the spatiotemporal Fourier transform of

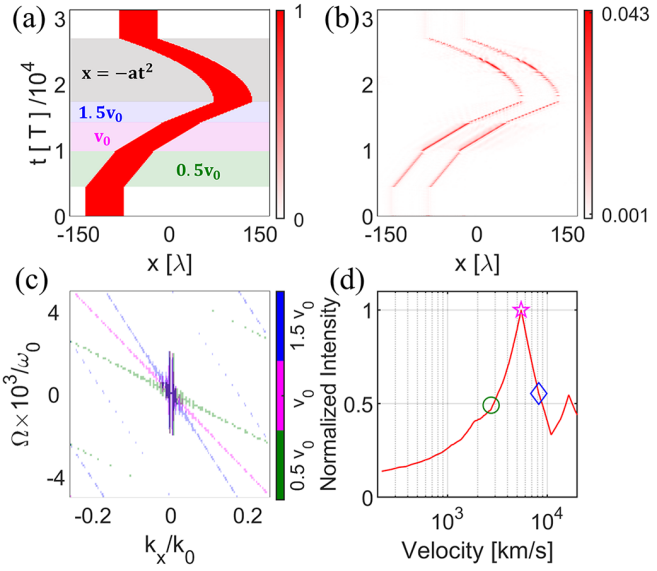

FIG. 4. Spatiotemporal edge detection of a moving object. (a) Time-dependent 1D image consisting of a segment with fixed width. The segment, initially at rest for a certain interval of time, starts moving with constant velocities 0.5 v_0 (green), v_0 (purple), and 1.5 v_0 (blue); then it negatively accelerates, and finally it stops. (b) Intensity of the output signal, highlighting the spatiotemporal edges. (c) Spatiotemporal Fourier transform of the signals corresponding to objects moving uniformly with three different speeds (see text for details). (d) Intensity of the spatiotemporal edges versus the speed of the segment.

a uniformly moving 1D shape, $F_v(k_x, \Omega)$, depends on the shape speed v. In Fig. 4(c) we consider an image [with the same size as in Fig. 4(a)] moving uniformly with velocities $v = 0.5\ v_0$ (green), $v = v_0$ (purple) and $v = 1.5\ v_0$ (blue). In order to highlight regions where each Fourier transform $F_v(k_x, \Omega)$ is largely different from zero, we associate to each speed v a binary function which is 0 when $|F_v(k_x, \Omega)|$ is smaller than 1.5% of its absolute maximum and 1 otherwise. By comparing Fig. 4(c) with Fig. 2(i), it becomes clear that the spatiotemporal Fourier transform $F_v(k_x, \Omega)$ of an image moving with velocity $v = v_0$ (purple color) optimally aligns with the high transmission lobes of the transfer function, thus leading to higher intensities of the spatiotemporal edges. This behavior is also summarized in Fig. 4(d), which shows the intensity of the spatiotemporal edges versus the image speed. As expected, the intensity peaks when $v_0 = \mathrm{cNA_t/NA_s} \approx 5800$ km/s (denoted by the purple star), and it decreases when the speed is either lower or higher than this optimal value.

We now discuss how the value of the characteristic speed v_0 impacts the output efficiencies and how it can be tailored. First, we emphasize that the speeds in Fig. 4 refer to the speed of an image projected onto the metasurface plane. In the case in which a far object is being imaged with a suitable imaging system, the speed of the image can be largely different than the actual speed of the physical object. Moreover, while in this particular device the speed v_0 that maximizes the output intensity is fairly high, we note that the dependence of the output intensity on the speed [Fig. 4(d)] is weak, and reasonable output intensities (∼20% of the peak) are still obtained for speeds that are over one order of magnitude smaller than v_0. Generally, for images moving at speeds $v < v_0$ and detected via the operation $\partial_t^m \partial_x^n$, the intensity of the spatiotemporal edges approximately scales as v^{2m}. Moreover, v_0 can be tailored through the MS nonlocality. The optimal speed v_0 is determined by $\mathrm{NA_t}$ and $\mathrm{NA_s}$, and it can be controlled by varying them. For a fixed temporal numerical aperture $\mathrm{NA_t}$, v_0 can be decreased by increasing the spatial numerical aperture $\mathrm{NA_s}$, which would also increase the spatial resolution. For the device considered here ($\mathrm{NA_s} = 0.26$), for example, the characteristic speed v_0 may be decreased by a factor of about 4 by working with larger $\mathrm{NA_s} \to 1$. Alternatively, for a fixed spatial numerical aperture, v_0 can be reduced by reducing the temporal numerical aperture $\mathrm{NA_t}$, linked to the spectral bandwidth of the MS performing the temporal differentiation [Figs. 2(b), 2(f)–2(g)]. In particular, v_0 can be largely reduced by increasing the quality factor of the MS. While the device considered here has a relatively low $Q \sim 60$, much higher Q factors can be achieved via, e.g., quasibound states in the continuum [50]. As an example, in [46] we show alternative designs with $v_0 \approx 3000$ and $v_0 \approx 1825$ km/s. Ultimately, a trade-off arises in the design process between bandwidth, speed and efficiency based on the highest speed present in the scene. More broadly, we emphasize that the event-based edge detection is not limited to moving physical objects, but it can be applied to images carried by pulses with extremely effective fast speeds [33], e.g., as a part of a neural network or neuromorphic computing platform.

Here, we have extended the concept of nonlocality engineering to the space-time domain, demonstrating a passive nonlocal metasurface that can perform analog mixed spatiotemporal differentiation on an input image. This functionality can be used to perform event-based edge detection, whereby the edges of an input image are enhanced only when the image intensity varies in time. Such event-based analog processing necessitates mixed spatiotemporal differentiation, and it cannot be achieved with other forms of spatiotemporal devices [35–38]. This nontrivial functionality—which may replace bulky circuit-based neuromorphic cameras—is achieved with a subwavelength silicon metasurface, whose design is compatible with conventional fabrication methods. We have shown that this MS can be used to detect the edges of an image whose intensity evolves in time, or to enhance the edges of a moving object. In this latter case, the intensity of the filtered image depends on the object speed, which opens interesting applications in the fields of sensing and vibration monitoring. Moreover, we have elucidated how the metasurface design can be tailored to maximally enhance objects

moving at desired speeds. The MS design can be readily extended to 2D image processing, paving the way towards optical analog event-based vision, leading to a vast spectrum of applications in fully passive, low-energy, and ultrafast optical computing.

Acknowledgments—This work was supported by the Air Force Office of Scientific Research and the Simons Foundation.

[1] J. Wang, L. Feng, W. Xue, and Z. Song, A survey on energy-efficient data management, SIGMOD Record **40**, 17 (2011).
[2] D. R. Solli and B. Jalali, Analog optical computing, Nat. Photonics **9**, 704 (2015).
[3] A. Silva, F. Monticone, G. Castaldi, V. Galdi, A. Alù, and N. Engheta, Performing mathematical operations with metamaterials, Science **343**, 160 (2014).
[4] A. Cordaro, B. Edwards, V. Nikkhah, A. Alù, N. Engheta, and A. Polman, Solving integral equations in free space with inverse-designed ultrathin optical metagratings, Nat. Nanotechnol. **18**, 365 (2023).
[5] F. Zangeneh-Nejad, D. L. Sounas, A. Alù, and R. Fleury, Analogue computing with metamaterials, Nat. Rev. Mater. **6**, 207 (2020).
[6] H. Caulfield and S. Dolev, Why future supercomputing requires optics, Nat. Photonics **4**, 261 (2010).
[7] V. Wiley and T. Lucas, Computer vision and image processing: A paper review, Int. J. Artif. Intell. Res. **2**, 1 (2018).
[8] C. D. Schuman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, Opportunities for neuromorphic computing algorithms and applications, NATO ASI series Series F, Computer and system sciences **2**, 10 (2022).
[9] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, Physics for neuromorphic computing, Nat. Rev. Phys. **2**, 499 (2020).
[10] X. Zhang and W. Dahu, Application of artificial intelligence algorithms in image processing, J. Visual Commun. Image Represent **61**, 42 (2019).
[11] W. Sun and C. Mo, High-speed real-time augmented reality tracking algorithm model of camera based on mixed feature points, J. Real-Time Image Proc. **18**, 249 (2021).
[12] R. Athale and D. Psaltis, Optical computing: past and future, Opt. Photonics News **27**, 32 (2016).
[13] R. S. Tucker, The role of optics in computing, Nat. Photonics **4**, 7 (2010).
[14] A. Cordaro, H. Kwon, D. Sounas, A. F. Koenderink, A. Alù, and A. Polman, High-index dielectric metasurfaces performing mathematical operations, Nano Lett. **19**, 8418 (2019).
[15] H. Kwon, A. Cordaro, D. Sounas, A. Polman, and A. Alù, Dual-polarization analog 2D image processing with nonlocal metasurfaces, ACS Photonics **7**, 1799 (2020).
[16] M. Cotrufo, A. Arora, S. Singh, and A. Alù, Dispersion engineered metasurfaces for broadband, high-NA, high-efficiency, dual-polarization analog image processing, Nat. Commun. **14**, 7078 (2023).
[17] Y. Zhou, H. Zheng, I. I. Kravchenko, and J. Valentine, Flat optics for image differentiation, Nat. Photonics **14**, 316 (2020).
[18] T. Zhu *et al.*, Generalized spatial differentiation from the spin Hall effect of light and its application in image processing of edge detection, Phys. Rev. Appl. **11**, 034043 (2019).
[19] T. Zhu, C. Guo, J. Huang, H. Wang, M. Orenstein, Z. Ruan, and S. Fan, Topological optical differentiator, Nat. Commun. **12**, 680 (2021).
[20] X. Ni, A. V. Kildishev, and V. M. Shalaev, Metasurface holograms for visible light, Nat. Commun. **4**, 1 (2013).
[21] S. Colburn, A. Zhan, and A. Majumdar, Metasurface optics for full-color computational imaging, Sci. Adv. **4**, eaar2114 (2018).
[22] N. A. Rubin, P. Chevalier, M. Juhl, M. Tamagnone, R. Chipman, and F. Capasso, Imaging polarimetry through metasurface polarization gratings, Opt. Express **30**, 9389 (2022).
[23] A. I. Kashapov, L. L. Doskolovich, E. A. Bezus, D. A. Bykov, and V. A. Soifer, Spatial differentiation of optical beams using a resonant metal-dielectric-metal structure, J. Opt. **23**, 023501 (2021).
[24] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, New York, 2019).
[25] J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, Greenwood Village, CO, 2005).
[26] H. Kwon, D. Sounas, A. Cordaro, A. Polman, and A. Alù, Nonlocal metasurfaces for optical signal processing, Phys. Rev. Lett. **121**, 173004 (2018).
[27] M. Cotrufo, S. Singh, A. Arora, A. Majewski, and A. Alù, Polarization imaging and edge detection with image-processing metasurfaces, Optica **10**, 1331 (2023).
[28] O. Y. Long, C. Guo, W. Jin, and S. Fan, Polarization-independent isotropic nonlocal metasurfaces with wavelength-controlled functionality, Phys. Rev. Appl. **17**, 024029 (2022).
[29] A. M. Weiner, Ultrafast optical pulse shaping: A tutorial review, Opt. Commun. **284**, 3669 (2011).
[30] A. M. Weiner, Femtosecond pulse shaping using spatial light modulators, Rev. Sci. Instrum. **71**, 1929 (2000).
[31] A. Momeni, K. Rouhi, and R. Fleury, Switchable and simultaneous spatiotemporal analog computing with computational graphene-based multilayers, Carbon **186**, 599 (2022).
[32] T. Knightley, A. Yakovlev, and V. Pacheco-Peña, Neural network design of multilayer metamaterial for temporal differentiation, Adv. Opt. Mater. **11**, 2202351 (2023).
[33] M. Cotrufo, S. Esfahani, D. Korobkin, and A. Alù, Temporal signal processing with nonlocal optical metasurfaces, arXiv:2403.09087.
[34] J. Sol, D. R. Smith, and P. del Hougne, Meta-programmable analog differentiator, Nat. Commun. **13**, 1713 (2022).
[35] J. Zhang, Q. Ying, and Z. Ruan, Time response of plasmonic spatial differentiators, Opt. Lett. **44**, 4511 (2019).
[36] C. Xu, Y. Wang, C. Zhang, B. Dagens, and X. Zhang, Optical spatiotemporal differentiator using a bilayer plasmonic grating, Opt. Lett. **46**, 4418 (2021).
[37] Y. Zhou, J. Zhan, R. Chen, W. Chen, Y. Wang, Y. Shao, and Y. Ma, Analogue optical spatiotemporal differentiator, Adv. Opt. Mater. **9**, 2002088 (2021).

[38] Y. Zhou *et al.*, Electromagnetic spatiotemporal differentiators, arXiv:2308.03797.
[39] L. L. Doskolovich, A. I. Kashapov, E. A. Bezus, and D. A. Bykov, Spatiotemporal optical differentiation and vortex generation with metal-dielectric-metal multilayers, Phys. Rev. A **106**, 033523 (2022).
[40] J. Huang, J. Zhang, T. Zhu, and Z. Ruan, Spatiotemporal differentiators generating optical vortices with transverse orbital angular momentum and detecting sharp change of pulse envelope, Laser Photonics Rev. **16**, 2100357 (2022).
[41] B. J. Shastri, A. N. Tait, T. Ferreira de Lima, W. H. P. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, Photonics for artificial intelligence and neuromorphic computing, Nat. Photonics **15**, 102 (2021).
[42] G. Gallego *et al.*, Event-based vision: A survey, IEEE Trans. Pattern Anal. Mach. Intell. **44**, 154 (2022).
[43] A. Lakshmi, A. Chakraborty, and C. S. Thakur, Neuromorphic vision: From sensors to event-based algorithms, WIREs Data Mining Knowl. Discovery **9**, e1310 (2019).
[44] K. Zhang, Y. Zhao, Z. Chu, and Y. Zhou, Event-based vision in magneto-optic Kerr effect microscopy, AIP Adv. **12**, 095315 (2022).
[45] A. Overvig, S. A. Mann, and A. Alù, Spatio-temporal coupled mode theory for nonlocal metasurfaces, Light Sci. Appl. **13**, 28 (2024).
[46] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.133.063801, which includes Ref. [47] for additional information on the design procedure, mathematical methods and numerical calculations.
[47] N. V. Golovastikov, D. A. Bykov, and L. L. Doskolovich, Spatiotemporal pulse shaping using resonant diffraction gratings, Opt. Lett. **40**, 3492 (2015).
[48] J. Huang, H. Zhang, B. Wu, T. Zhu, and Z. Ruan, Topologically protected generation of spatiotemporal optical vortices with nonlocal spatial mirror symmetry breaking metasurface, Phys. Rev. B **108**, 104106 (2023).
[49] H. Zhang, Y. Sun, J. Huang, B. Wu, Z. Yang, K. Y. Bliokh, and Z. Ruan, Topologically crafted spatiotemporal vortices in acoustics, Nat. Commun. **14**, 6238 (2023).
[50] K. Koshelev, S. Lepeshov, M. Liu, A. Bogdanov, and Y. Kivshar, Asymmetric metasurfaces with high-Q resonances governed by bound states in the continuum, Phys. Rev. Lett. **121**, 193903 (2018).