



Single-shot real-time sub-nanosecond electron imaging aided by compressed sensing: Analytical modeling and simulation

Xianglei Liu^a, Shian Zhang^b, Aycan Yurtsever^{a,*}, Jinyang Liang^{a,*}

^a Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, 1650 boulevard Lionel-Boulet, Varennes, Québec, J3X1S2, Canada

^b State Key Laboratory of Precision Spectroscopy, East China Normal University, 3663 North Zhongshan Road, Shanghai, 200062, PR China

ARTICLE INFO

Keywords:

Transmission electron microscopy
Ultrafast imaging
Compressed sensing
Streak imaging

ABSTRACT

Bringing ultrafast (nanosecond and below) temporal resolution to transmission electron microscopy (TEM) has historically been challenging. Despite significant recent progress in this direction, it remains difficult to achieve sub-nanosecond temporal resolution with a single electron pulse, in real-time (i.e., duration in which the event occurs) imaging. To address this limitation, here, we propose a methodology that combines laser-assisted TEM with computational imaging methodologies based on compressed sensing (CS). In this technique, a two-dimensional (2D) transient event [i.e. (x, y) frames that vary in time] is recorded through a CS paradigm, which consists of spatial encoding, temporal shearing via streaking, and spatiotemporal integration of an electron pulse. The 2D image generated on a camera is used to reconstruct the datacube of the ultrafast event, with two spatial and one temporal dimensions, via a CS-based image reconstruction algorithm. Using numerical simulation, we find that the reconstructed results are in good agreement with the ground truth, which demonstrates the applicability of CS-based computational imaging methodologies to laser-assisted TEM. Our proposed method, complementing the existing ultrafast stroboscopic and nanosecond single-shot techniques, opens up the possibility for single-shot, real-time, spatiotemporal imaging of irreversible structural phenomena with sub-nanosecond temporal resolution.

1. Introduction

Imaging materials with combined high spatial and temporal resolutions merits our attention to understanding the origins of many microscopic phenomena. However, developing such spatiotemporal imaging techniques has been traditionally challenging because of the limitations of the elementary particles used in the imaging processes. Photons and electrons, two fundamental particles that account for the most, if not all, imaging technologies, are dissimilar in terms of the spatiotemporal domains they can access. Photons, with no dispersion in free space and with no elementary charge, excel in forming ultrashort pulses that can propagate long distances and, hence, can be used for extremely high (up to attosecond) temporal studies. Fast electrons, on the other hand, with their picometer wavelengths and strong interaction cross-sections, excel in forming images with the highest spatial resolution (sub-angstrom) available today.

Methodologies based on the former, such as ultrafast optical spectroscopy, have offered access to the first “moments” of fundamental phenomena, including the birth of chemical species and evolution of molecular bonds (Zewail, 1990, 2000), phonon dynamics (Cavalleri

et al., 2004; Elsayed-Ali et al., 1987), and valence-band electronic transitions (Sjodin et al., 1998). While excelling in temporal resolution, photons lack the extreme spatial resolution that electrons can achieve in transmission electron microscopy (TEM).

In the last couple of decades, it was possible to make photons and electrons work together in one microscope (Bostanjoglo, 2002; Bostanjoglo et al., 1991; Bostanjoglo and Weingärtner, 1997; Bücker et al., 2016; Feist et al., 2017; Kim et al., 2008; LaGrange et al., 2006; Meurig, 2009; Musumeci et al., 2017; Yurtsever et al., 2012) to achieve combined spatiotemporal resolutions that were not possible before. At the California Institute of Technology, a TEM system was modified to merge it with a femtosecond laser (Meurig, 2009). In such an ultrafast TEM, instead of using continuous electron beams, ultrashort electron pulses created by the photoelectric effect were used for imaging, while optical pulses were used to clock the sample synchronously. With such a pump-probe scheme, it was possible to achieve few-nanometers real-space resolution combined with sub-picosecond temporal resolution stroboscopically (Yurtsever et al., 2012), which represents the current resolution limits in spatiotemporal imaging. Complementarily, another photon-electron imaging methodology was developed at the Lawrence

* Corresponding authors.

E-mail addresses: aycan.yurtsever@emt.inrs.ca (A. Yurtsever), jinyang.liang@emt.inrs.ca (J. Liang).

<https://doi.org/10.1016/j.micron.2018.11.003>

Received 3 October 2018; Received in revised form 15 November 2018; Accepted 15 November 2018

Available online 16 November 2018

0968-4328/ © 2018 Elsevier Ltd. All rights reserved.

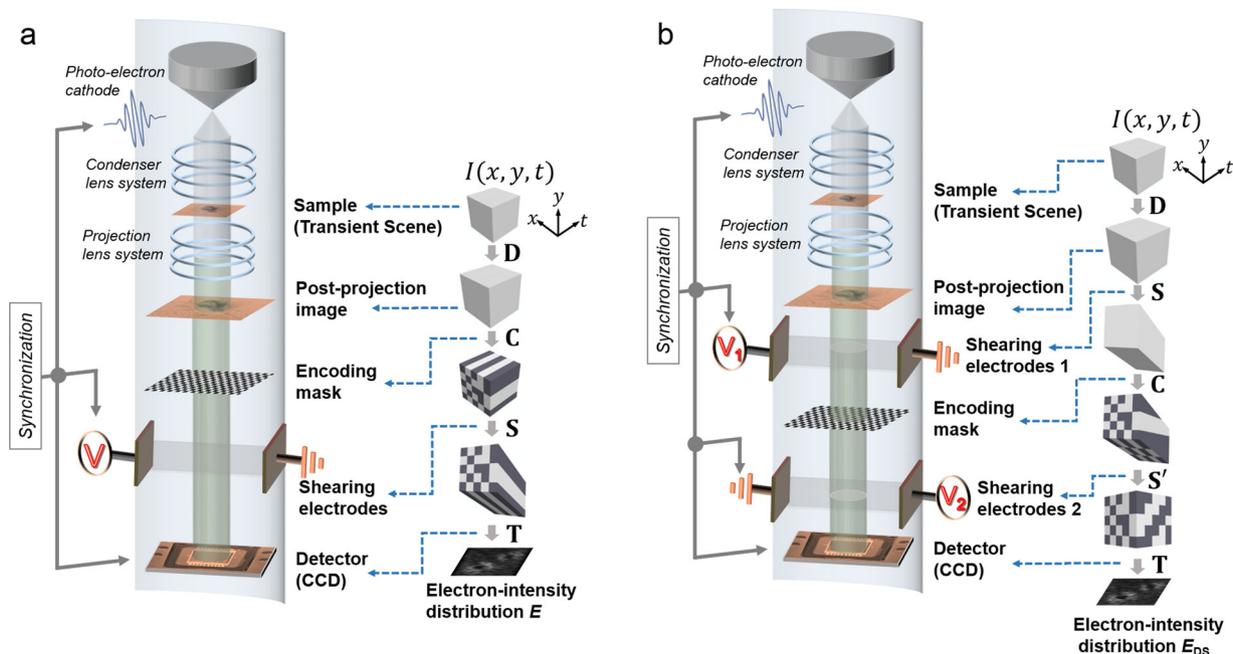


Fig. 1. Proposed methods for single-shot real-time ultrafast electron imaging. (a), Hardware schematic of CUTEM and its forward model. C , spatial encoding operator; D , electron spatiotemporal distortion operator; S , temporal shearing operator; T , spatiotemporal integration operator. (b), Hardware schematic of DS-CUTEM and its forward model. S' , reverse temporal shearing operator. t , time; x, y , spatial coordinates.

Livermore National Laboratory (LLNL) (LaGrange et al., 2012), following the original work of O. Bostanjoglo and colleagues (Bostanjoglo, 2002; Bostanjoglo et al., 1991; Bostanjoglo and Weingärtner, 1997). This technique was based on using several intense electron pulses to image a photo-excited sample, and then deflecting these pulses successively to fill a CCD camera with several frames. This single-shot, real-time (i.e., duration in which the event occurs), ultrafast imaging approach enabled visualizing irreversible phenomena, albeit with limited sequence depth (i.e., the maximum number of frames that can be captured in one acquisition), real-space resolution of tens of nanometers, and temporal resolution in the nanosecond range (LaGrange et al., 2012).

Recently, compressed sensing (CS)—a scheme that allows underdetermined reconstruction of sparse scenes (Hunt et al., 2013)—has been introduced into non-laser-assisted TEM (Stevens et al., 2015). In this proposal, a coded-aperture-based CS method (Llull et al., 2013) was implemented to increase the TEM's frame rate. After the sample, the spatiotemporally modulated electron beam was encoded by a binary transmissive mask loaded on a moving piezoelectric stage. Different encoding masks were attached to individual temporal frames. These spatially encoded frames were integrated in a single camera exposure during the acquisition process. Their simulation showed that a movie could be reconstructed via CS inversion by using the captured single image. This proposed scheme would allow increasing TEM's frame rate to several kilohertz (Llull et al., 2013). However, constrained by the translational speed of the piezoelectric stage, the limited frame rate prevents this scheme from achieving sub-nanosecond temporal resolution. CS has also been implemented in the time domain with a TEM system (Reed, 2017; Reed et al., 2017). In this technique, four replicated images (arranged as a 2×2 array) of the dynamic scene were generated on an image sensor by a high-speed deflector. A dynamic shutter, encoded by a distinct grayscale code sequence for each image, modulated the temporal integration process. The acquired data were fed into a CS-based reconstruction algorithm to generate 10 frames at the kilohertz level. Thus, this method has limitations in the imaging speed and the sequence depth.

To overcome these limitations, here, we introduce a new methodology in laser-assisted TEM, inspired by compressed ultrafast

photography (CUP) that has demonstrated breakthroughs in recording optical ultrafast transient scenes (Gao et al., 2014; Liang et al., 2017, 2018). CUP, based on streak imaging and CS, has exhibited the fastest receive-only imaging modality at up to 10 trillion frames per second with a 0.58-ps temporal resolution. Leveraging a CS-based data acquisition and image reconstruction paradigm, CUP adds another spatial dimension into the conventional streak camera, enabling two-dimensional (2D) ultrafast (i.e., x, y, t) imaging with one camera exposure (Liang and Wang, 2018). CUP has enabled capturing, for the first time, a number of transient optical phenomena, such as real-time fluorescence emission (Gao et al., 2014), propagation of a scattering-induced photonic Mach cone (Liang et al., 2017), and temporal focusing of a femtosecond laser pulse (Liang et al., 2018). Thus far, CUP has been applied only in optical imaging. However, CUP's principle is generic, indicating the possibility of extending it to imaging with electrons.

In this work, we propose two single-shot real-time ultrafast electron imaging configurations, synergizing CUP and laser-assisted TEM. The first configuration, termed compressed ultrafast TEM (CUTEM), can be integrated into a laser-assisted TEM system without excessive modifications. The second configuration, termed dual-shearing (DS)-CUTEM, can offer improved technical specifications, but with more elaborate modifications to the hardware. For each configuration, we examined how each manipulation of a transient scene during electron propagation and data acquisition affected reconstructed image quality by using experimental data as the ground truth. These quantitative analyses guided us to optimize the reconstructed imaging quality. As a result, we were able to use the proposed configurations to reconstruct spatiotemporal datacubes, which were found to be in good agreement with the ground truth. This study verifies the feasibility of implementing CS-aided ultrafast imaging modalities to TEM. The proposed schemes have potential to achieve single-shot ultrafast imaging in real time with sub-nanosecond temporal resolution and nanometers real-space resolution.

The rest of the paper is organized as follows. In Section 2, we describe the system designs with the analytical models of associated data acquisition and imaging reconstruction. In Section 3, we present the optimization of reconstructed image quality and proof-of-concept demonstrations using numerical simulation. In Section 4, we discuss how

the proposed techniques could guide future experiments, followed by a short summary that concludes this paper.

2. Method

2.1. CUTEM

The proposed CUTEM [Fig. 1(a)] technique can be experimentally realized by integrating a mask and shearing electrodes to a laser-assisted TEM machine. In CUTEM, a single nanosecond electron pulse probes a transient event initiated in a sample to generate a spatiotemporally modulated electron density function, $I(x, y, t)$. The transmitted electron pulse, then, experiences several manipulations, each of which is accounted by an operator (detailed in the following paragraph). Specifically, the pulse undergoes spatiotemporal electron distortion during the ensuing propagation within the TEM system. A binary transmission mask, with a specifically designed pattern, is placed on the path of the imaging electrons to encode $I(x, y, t)$. Produced by existing nanofabrication tools, this mask blocks the electrons at certain regions in space and allows them to pass through at others. Following the mask, a time-varying voltage, applied to a pair of electrodes, temporally shears $I(x, y, t)$. Finally, a single streak image, denoted as E , is formed on a CCD camera via spatiotemporal integration (spatially integrating over each CCD pixel and temporally integrating over the exposure time). The mask and the shearing electrodes can be inserted to the TEM machine using one of the ports after the projection lens system.

Mathematically, the above-described data acquisition process can be expressed as a forward model [Fig. 1(a)] by

$$E = \mathbf{TSCD}I(x, y, t) \quad (1)$$

where the linear operator \mathbf{D} accounts for electron's spatiotemporal distortion induced by all experimental factors. \mathbf{C} represents spatial encoding by the binary transmission mask, \mathbf{S} represents temporal shearing by the time-varying voltage applied to the electrodes, and \mathbf{T} represents spatiotemporal integration by the CCD camera. For simplicity in denotation, we use a linear operator \mathbf{O} to represent the entire data acquisition process (i.e., $\mathbf{O} = \mathbf{TSCD}$). After data acquisition, we retrieve the transient scene computationally. Given the known operator \mathbf{O} and the spatiotemporal sparsity of the transient scene, $I(x, y, t)$ can be recovered by solving the inverse problem of Eq. (1). This process can be formulated as

$$\hat{I} = \underset{I}{\operatorname{argmin}}\{\|E - \mathbf{O}I\|_2^2 + \lambda\Phi(I)\} \quad (2)$$

where λ is a weighting parameter, and $\Phi(\cdot)$ is the regularization function. In Eq. (2), the minimization of the first term, $\|E - \mathbf{O}I\|_2^2$, occurs when the actual measurement E closely matches the estimated solution $\mathbf{O}I$, while the minimization of the second term, $\Phi(I)$, encourages I to be piecewise constant (i.e., sparse in the spatial gradient domain). The weighting of these two terms is adjusted by λ to lead to the results that are most consistent with the ground truth. Given the number of rows and columns of the CCD camera to be N_r and N_c , the number of voxels in the reconstructed datacube (i.e., N_x , N_y , and N_t) must meet the requirement of

$$\begin{aligned} N_x &\leq N_c \quad \text{and} \\ N_y + N_t - 1 &\leq N_r \end{aligned} \quad (3)$$

2.2. Dual-shearing (DS)-CUTEM

To increase the size of reconstructed datacube, we further propose a variant referred to as dual shearing (DS)-CUTEM, whose system schematic is shown in Fig. 1(b). The difference of DS-CUTEM with respect to CUTEM is that $I(x, y, t)$ is sheared twice in opposite directions by two pairs of electrodes, which sandwich the spatial encoding mask. A

similar voltage ramping waveform is used to control both pairs to avoid asymmetric shearing. The delay of shearing onsets is determined by the distance between these two deflectors and the electron pulse's propagation speed. The streak image recorded on the CCD camera is denoted as E_{DS} .

Mathematically, the forward model of DS-CUTEM can be expressed as

$$E_{\text{DS}} = \mathbf{TS}'\mathbf{CSD}I(x, y, t) \quad (4)$$

where \mathbf{S}' represents temporal shearing in a reverse direction to that of \mathbf{S} . After data acquisition, DS-CUTEM uses the similar computational reconstruction method as CUTEM to retrieve the result, expressed by

$$\hat{I} = \underset{I}{\operatorname{argmin}}\{\|E_{\text{DS}} - \mathbf{O}'I\|_2^2 + \lambda\Phi(I)\} \quad (5)$$

where the linear operator $\mathbf{O}' = \mathbf{TS}'\mathbf{CSD}$.

Compared with CUTEM, the dual-shearing operation implemented in DS-CUTEM allows the temporal course of a specific spatial position in the scene to be integrated at the same pixel of the CCD camera, circumventing the requirement [i.e., Eq. (3)] that is imposed by CUTEM. Consequently, each frame in the reconstructed datacube could have a maximum size of $N_x \times N_y = N_c \times N_r$, and N_t would be limited by the dynamic range of the CCD camera. In addition, the dual-shearing operation implemented in DS-CUTEM enables designing a more incoherent measurement matrix. In particular, while encoded by the same mask in CUTEM, each frame in $I(x, y, t)$ is imprinted by a different random binary mask in DS-CUTEM due to the first shearing operation prior to the spatial encoding. Consequently, DS-CUTEM has less mutual coherence between the measurement matrix and the sparsity basis of the transient scene than that of CUTEM, which yields an improved reconstructed image quality (detailed in Section 3).

2.3. Comparison of CUTEM and DS-CUTEM with framing TEM

The proposed two configurations could exceed framing TEM [i.e. where the CCD camera is filled with spatially-separated frames, like in the LLNL TEM system (LaGrange et al., 2012)] in the temporal resolution and the sequence depth. Both CUTEM and DS-CUTEM use spatial encoding and temporal shearing to tag each frame with a spatiotemporal "barcode". This prior information, along with the sparsity in the transient scene, allows spatiotemporal mixing of adjacent frames, which can be recovered in image reconstruction. The reconstructed frame rate, r , is thus determined by

$$r = v/d \quad (6)$$

where v is the temporal shearing velocity, and d is the CCD's pixel size along the temporal shearing direction. As a result, for a given recording time window, the proposed configurations can largely improve the temporal resolution, compared with the framing TEM. For instance, using a high-speed electrostatic deflector, the framing TEM (LaGrange et al., 2012) can acquire up to 5×5 frames [Fig. 2(a)], each of which has 400×400 pixels in size to fill up a sensor with pixels of $N_c \times N_r = 2000 \times 2000$. The inter-frame time is 25 ns. In comparison, spatiotemporal mixing allowed in CUTEM could produce an inter-frame time of 62.5 ps. With the consideration of temporal blurring effect (Gao et al., 2014), the estimated temporal resolution is approximately 0.3 ns. According to Eq. (3), in CUTEM, the x axis is used only for recording the spatial information, which allows making full use of the N_c (i.e., $N_x = 2000$). In the y axis, the pixels are distributed to record both spatial and temporal information. Physically, this restriction means that the size of captured 2D image has to be less than the full size of the sensor [Fig. 2(b)]. Therefore, if we keep the number of pixels in the y axis the same as the framing TEM (i.e., $N_y = 400$), the maximum sequence depth is around $N_t = 1600$.

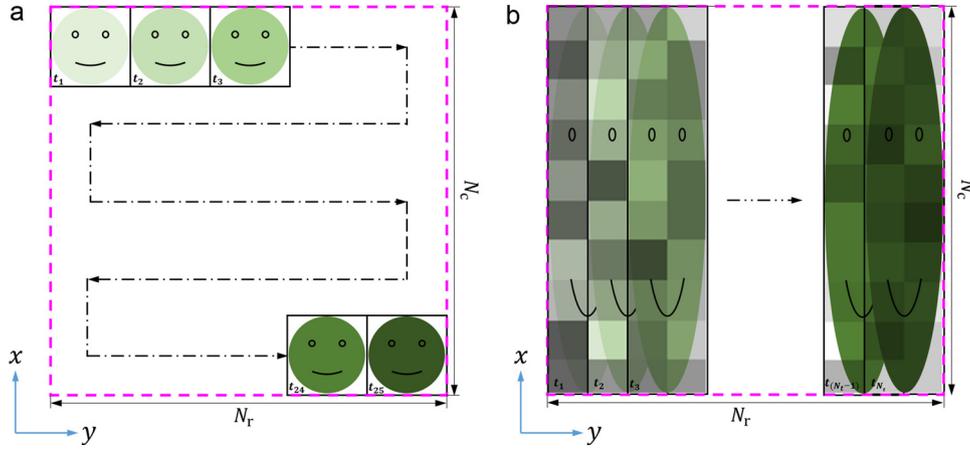


Fig. 2. Illustrative comparison of the captured images on the camera by (a) framing TEM and (b) CUTEM. The dashed box indicates the camera's sensor area. The dash-dotted arrow lines in (a) and (b) indicate the direction of raster scanning of framing TEM and the temporal shearing direction of CUTEM, respectively. N_r and N_c , number of rows and columns of the camera. In (b), the adjacent two frames can be offset by as few as one pixel.

3. Results

3.1. Simulation setup

To test the proposed methods, two transient scenes were used. The first scene [Fig. 3(a)] depicts near-field intensity dynamics of dipoles of two close-by silver nanoparticles excited by an ultrashort laser pulse (referred to hereafter as the ‘dipole scene’) (Yurtsever and Zewail, 2012). While holding fixed spatial positions, the dipoles have fluctuating intensities in time. Second, as shown in Fig. 3(b), a transient scene of a moving nanoscale cantilever was used (referred to as the ‘cantilever scene’ hereafter) (Flannigan et al., 2009). Complementary to the dipole scene, the cantilever moves in space, while its intensity is almost constant. Both datasets are experimentally acquired and previously published (Flannigan et al., 2009; Yurtsever and Zewail, 2012).

The general framework of our simulation is as follows. First, the sizes of both the dipole scene and cantilever scene were set to $N_x \times N_y \times N_t = 300 \times 300 \times 30$. Then, both scenes were used for the forward models of the two proposed methods [i.e., Eqs. (1) and (4)] to generate E and E_{DS} , respectively. To maximally mimic the experimental conditions, no noise filtering was conducted to any frames in either scene. The sizes of E and E_{DS} were $N_x \times N_{y,t} = 300 \times 329$ and $N_x \times N_{y,t} = 300 \times 300$, respectively. Here, $N_{y,t}$ represents the number of pixels in the axis where the spatial and temporal information are mixed. Finally, E and E_{DS} were fed into the two-step iterative shrinkage/thresholding (TwIST) algorithm (Bioucas-Dias and Figueiredo, 2007) to retrieve $I(x, y, t)$ by solving Eqs. (2) and (5). Total variation (TV) was used as the regularization function (Vogel and Oman, 1996). Compared with other regularization functions, such as l_0 or l_1 norms, TV regularization has superior performance in denoising while preserving important details in images (Strong and Chan, 2003).

3.2. Examination of various operators and reconstructed image quality

To investigate how each operator impacts the reconstructed image quality, we used different combinations of encoding masks, encoding pixel sizes (defined as how many pixels of the sensor each encoding pixel contains), and signal-to-noise ratios (SNRs) in both proposed methods. First, the encoding mask provides the necessary prior information to extract the temporal information in image reconstruction. Following the established theories (Elad, 2007; Koller et al., 2015; Tsiliigianni et al., 2014), a pseudo-random binary mask [shown in Fig. 3(c) and referred to hereafter as the ‘conventional mask’] was implemented. This mask was constructed by binarizing a random matrix whose element values follow a Gaussian distribution. While

incoherent with the most sparse representation of the scene, the conventional mask might not be the optimal choice for a given sparse basis. To optimize the encoding mask, we used a column-normalized random matrix \tilde{C} to generate its Gram matrix (Bernstein, 2005), defined by $\tilde{C}^T \tilde{C}$. When the Gram matrix approaches a unit matrix H , the mutual coherence reaches the minimum. Mathematically, this optimization process can be formulated as

$$\hat{C} = \underset{\tilde{C}}{\operatorname{argmin}} \|\tilde{C}^T \tilde{C} - H\|_F^2 \quad (7)$$

where $\|\cdot\|_F^2$ represents Frobenius norm. In practice, a gradient descent method (Abolghasemi et al., 2010, 2012) was implemented to generate the optimized random binary mask [Fig. 3(d)]. For simplicity, we denote this mask as the ‘optimized mask’.

Besides the mask pattern, we also optimized the encoding pixel size. On one hand, to satisfy the Nyquist sampling theorem, the smallest feature in both spatial and temporal domains has to be sampled by at least two encoding pixels, meaning that the smaller pixel size aids spatial and temporal resolutions. On the other hand, the smaller pixel size is more vulnerable to noise in the measurement and thus demands higher SNRs for transient scenes (Kellman and McVeigh, 2005). Thus, the optimized encoding pixel size can be chosen by balancing the achievable resolution and noise.

Finally, the SNRs in E and E_{DS} , associated with the operator \mathbf{T} [Eqs. (1) and (4)], were analyzed. Here, we defined the SNR as the quotient of the maximal intensity to the standard deviation of the noise in the acquired images (i.e., E and E_{DS}). To vary the SNR, we added zero-mean Gaussian noise with variances from 0 to 0.5.

To quantitatively compare the reconstructed image quality, the structural similarity (SSIM) (Wang et al., 2004) was used as the merit function. SSIM is a perception-based assessment model that considers image degradation as perceived change in structural information while also incorporating significant perceptual phenomena, including both luminance and contrast terms. Compared with conventional image quality assessment methods, such as mean squared error and peak signal-to-noise ratio, SSIM has a better capability to represent the perceived visual quality.

Fig. 3(e) and (f) show SSIM values of the reconstructed dipole and cantilever scenes, using both the conventional and optimized encoding masks and with different encoding pixel sizes, for both CUTEM and DS-CUTEM. For both scenes, dual-shearing operation produces better reconstruction image quality than its single-shearing counterpart because of its more incoherent measurement matrix. Fig. 3(e) and (f) also illustrate that mask optimization improves reconstructed image quality. Although the mask optimization was conducted only in the spatial

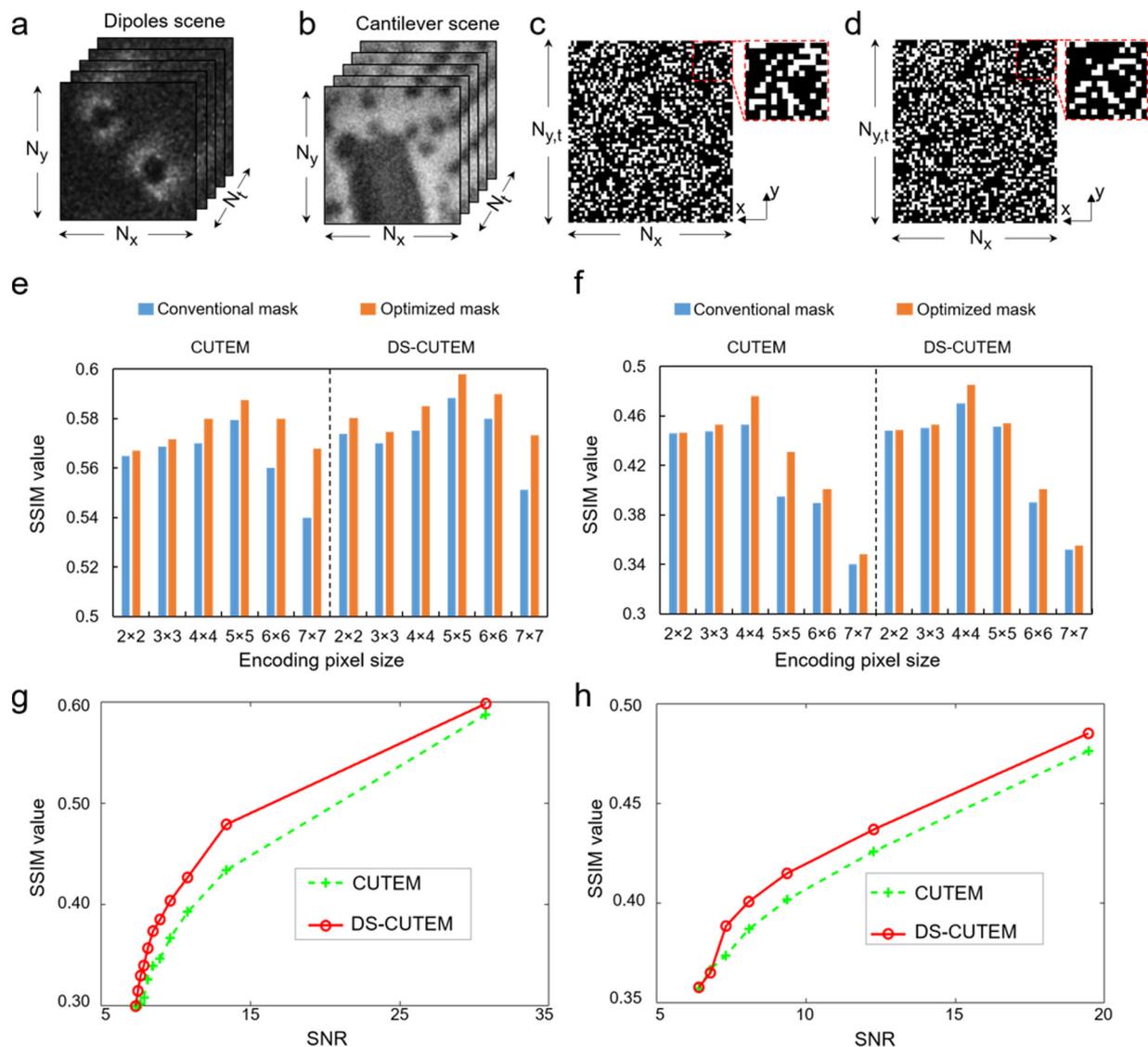


Fig. 3. Examination of the reconstructed image quality to the types of encoding masks, encoding pixel sizes, and signal-to-noise ratios (SNRs). (a), Dipole scene. (b), Cantilever scene. Part of the content in (a) and (b) was adapted from (Flannigan et al., 2009; Yurtsever and Zewail, 2012) with permissions. (c), Conventional binary mask. (d), Optimized binary mask. Insets in (c) and (d) show the details of local features of the encoding masks; x, y , spatial coordinates. (e), SSIM values of the reconstructed dipole scene versus different encoding pixel sizes and the two types of encoding masks for both proposed methods. (f), As (e), but shows the results using the cantilever scene. (g), SSIM values with the respect to SNRs for the dipole scene (using 5×5 encoding pixel size). (h), As (g), but shows the SSIM-SNR relations for the cantilever scene (using 4×4 encoding pixel size). The optimized binary masks were used for obtaining the results in (g) and (h).

domain, because C is a part of measurement matrices O and O' , this operation still assisted in minimizing the mutual coherence between these measurement matrices and the sparsity basis, leading to a better reconstruction image quality. It is worth noting that although only demonstrated in the spatial domain in this work, mask optimization can be implemented for other sparsity bases, including discrete cosine transformation (Lustig et al., 2007) and discrete wavelet transformation (Lustig et al., 2008). Finally, we found the optimal encoding pixel size to be 5×5 for the dipole scene and 4×4 for the cantilever scene, respectively. The varied optimal sizes are ascribed to the different imaging contents and characteristics in the two scenes.

Fig. 3(g) and (h) present the changes in SSIM values of the dipole scene and the cantilever scene versus the SNRs of E and E_{DS} . The optimized encoding patterns and the optimized encoding pixel sizes were used for both scenes in this quantification. With decreasing SNRs, although the reconstructed image quality descends drastically for both proposed methods, DS-CUTEM performs consistently better than CUTEM until the SNR approaches five. There, the SSIM values for both

proposed methods converge and remain relatively unchanged for lower SNRs, because the noise has dominated the measurement results.

Besides the above-described optimization, we accounted for the influence of the space-charge effect on image distortion by introducing the operator \mathbf{D} in Eqs. (1) and (4). The space-charge effect degrades both spatial resolution and temporal resolution. First, by distorting the probing electron pulse, it causes deformation of the system's point spread function. The distortion would lead to crosstalk (i.e., electrons that are supposed to be recorded by a certain pixel on the image sensor are actually recorded by a different pixel) in the spatiotemporal integration. Because spatial information and temporal information are mixed during the data acquisition in both CUTEM and DS-CUTEM, this crosstalk would affect the reconstruction accuracy in both the spatial domain and the time domain. Finally, this inaccurate reconstruction will transfer to the degraded spatial and temporal resolutions.

We modeled electrons' spatial and temporal expansion by a spatiotemporal Gaussian model. The expansion ratio in the model is defined as the quotient between the size of the distorted transient scene

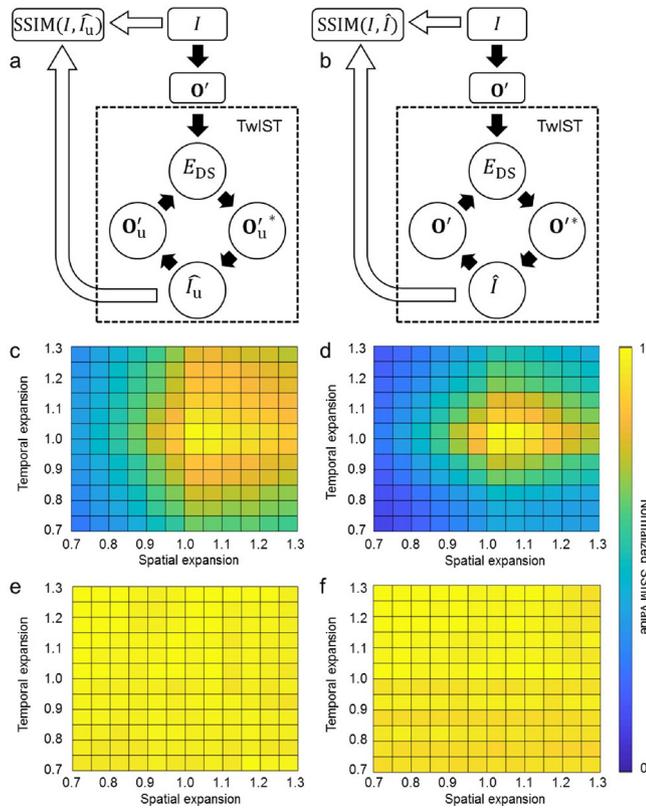


Fig. 4. Comparison between the distortion and correction in DS-CUTEM. (a), Flowchart of reconstruction without distortion correction. (b), Flowchart of reconstruction with distortion correction. (c), Normalized SSIM values of reconstructed datacubes of the dipole scene without distortion correction. (d), As (c), but shows results for the cantilever scene. (e), Normalized SSIM values of reconstructed datacubes of dipole scene with distortion correction. (f), As (e), but shows results for the cantilever scene.

and that of the original one. A ratio smaller than one means spatial or temporal shrinkage. Both proposed methods were implemented to account for \mathbf{D} and produced comparable results. For simplicity, only the results of DS-CUTEM are shown. The flow chart of the reconstruction process with electron distortion is shown in the Fig. 4(a). The transient scene first went through the operator \mathbf{O}' to generate E_{DS} [see Eq. (4)]. Then, E_{DS} was used as an input into the TwiST algorithm to retrieve the estimated measurement $\hat{I}_u(x, y, t)$. The operator \mathbf{O}'_u represents \mathbf{TSCS} , and its inverse operator is represented as \mathbf{O}'_{u^*} . Here, the subscript “u” stands for “uncorrected”. Finally, we used the estimated measurement $\hat{I}_u(x, y, t)$ and input transient scene $I(x, y, t)$ to calculate the SSIM value. As a comparison, the flow chart of the reconstruction process with distortion correction is shown in Fig. 4(b). \mathbf{O}'' denotes the inverse operator of \mathbf{O}' .

Fig. 4(c) shows the normalized SSIM values of reconstructed dipole scene with spatial and temporal expansions, both with ratios between 0.7 and 1.3. The expansions in both spatial and temporal domains deteriorated the reconstructed image quality, manifesting in the radially decreasing SSIM values from the origin (i.e., no spatial or temporal expansion). Similar degradation trend is observed for the cantilever scene [Fig. 4(d)]. The result after correcting for electrons’ spatio-temporal distortion through the operator \mathbf{D} is shown in Fig. 4(e) and (f). In both cases, corrections significantly alleviate the degradation, demonstrated as the high normalized SSIM values (between 0.95 and 0.96) across various spatiotemporal expansion.

It is worth pointing out that although we used a simple model for the operator \mathbf{D} to compensate for the spatiotemporal distortions of the electron pulse, the reconstruction approach presented here is universal: as long as the operator \mathbf{D} is known (e.g., obtained through a point-by-

point calibration), image distortions due to electron repulsions can be accounted for. In a TEM system, the electron pulse can propagate a significant distance, with several crossovers in the imaging/projection lens system. For more precise reconstruction, these effects should be accounted for via elaborate studies of the electron trajectories. Nevertheless, for moderate electron densities in a pulse, the Gaussian model used here, as a first order approximation to model, should be sufficient.

3.3. Feasibility demonstration

We leveraged the proposed methodology in reconstructing the two transient scenes. The represented frames of the dipole scene (as the ground truth) are shown in Fig. 5(a). The corresponding frames reconstructed by CUTEM and DS-CUTEM are presented in Fig. 5(b) and (c), respectively. A frame-to-frame comparison of the two reconstructed datacubes with the ground truth is shown in Media 1. Both CUTEM and DS-CUTEM significantly removed the noises presented in the dipole scene, demonstrated by a clean background in each reconstructed frame. Fig. 5(d) illustrates the change in normalized average intensity across all frames. The reconstruction results of CUTEM and DS-CUTEM are consistent with the ground truth. For a more detailed analysis, in Fig. 5(e), we plot the intensity profiles of a diagonal line in the ground truth and the reconstructed results [marked by the red solid line in Frame 15 in Fig. 5(a)–(c)]. The reconstruction by both CUTEM and DS-CUTEM well complies with the ground truth.

Six representative frames of the cantilever scene (as the ground truth) and its reconstructed images are shown in Fig. 6(a)–(c). The reconstructed movies are shown in Media 2. Similar to the dipole scene, the noise reduction is clearly represented in Fig. 6(b) and (c), and the reconstruction is in good agreement with the ground truth [Fig. 6(d)]. Fig. 6(e) shows the intensity profiles of a selected horizontal line in the ground truth and reconstructed results [marked by the solid line in Frame 5 in Fig. 6(a)–(c)], which again demonstrates that the good agreement between the reconstructed results with the ground truth.

4. Discussion and conclusions

Our analytical models and simulation have shown that the single-shot ultrafast electronic imaging with sub-nanosecond temporal resolution could be realized by integrating CS-aided ultrafast imaging modalities to laser-assisted TEM. In principle, through hardware modifications detailed in Fig. 1, CUTEM and DS-CUTEM will be able to record a single spatiotemporally modulated image in measurement and time-resolved frames through reconstruction. Moreover, compared with the previous methods (Reed, 2017; Stevens et al., 2015), both proposed configurations in our work encode the transient scene with a static mask. Consequently, no mechanical movement will be induced during the image acquisition, which would improve the repeatability and accuracy in calibration and experiments. In practice, to generate the image with high SNRs, nanosecond pulsed electron beams with more than 10 million electrons will be utilized. Such electron pulses can be generated using the photoelectric effect and powerful lasers. Temporal shearing and image recording needed for CUTEM and DS-CUTEM can be achieved with electronics with a rise time in the order of hundreds of picoseconds to nanoseconds (LaGrange et al., 2012). At Institut National de la Recherche Scientifique (INRS), a TEM system using these technologies is under construction.

In summary, we have presented the design and simulation studies of CUTEM and DS-CUTEM—two CS-aided single-shot real-time ultrafast electron imaging methods. Transient scenes experimentally captured by laser-assisted TEM were used to test the fidelity of the proposed methods. Different types of mask patterns and encoding pixel sizes were examined to optimize reconstructed image quality for the given SNR. The reconstruction results with both electronic distortion and corresponding correction processes were also provided. Based on these

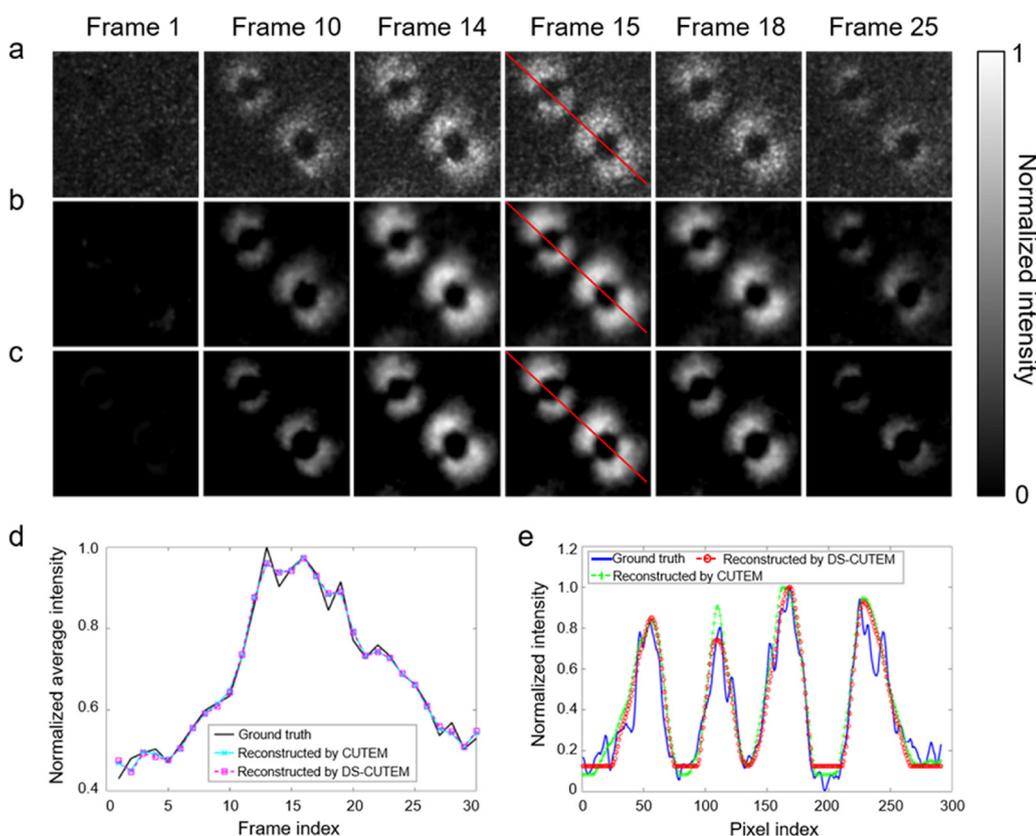


Fig. 5. Reconstruction of the dipole scene using CUTEM and DS-CUTEM. (a), Representative frames of dipole scene as the ground truth [adapted from (Yurtsever and Zewail, 2012) with permission]. (b) and (c), Corresponding representative frames of reconstruction using CUTEM and DS-CUTEM, respectively. (d), Normalized average intensity across all frames of dipole scene and the two reconstructed results. (e), Comparison of normalized intensity profiles of a diagonal line in the dipole scene and the reconstruction results [marked by the solid line in Frame 15 in (a)–(c)]. A frame-by-frame comparison of the two reconstructed datacubes with the ground truth is shown in Media 1.

investigations, high-quality reconstruction of both transient scenes have been demonstrated in simulation. Our future work will be to apply both proposed methods in the advanced TEM infrastructure at INRS for imaging non-repetitive ultrafast events, such as structure dynamics of

superheated nanoparticles (Gorkhover et al., 2016) and phase transition in crystals (Baum et al., 2007; Gedik et al., 2007) at sub-nanosecond temporal resolution and nanometer real-space resolution. In addition, we plan to graft the concept of computational tomography (CT)

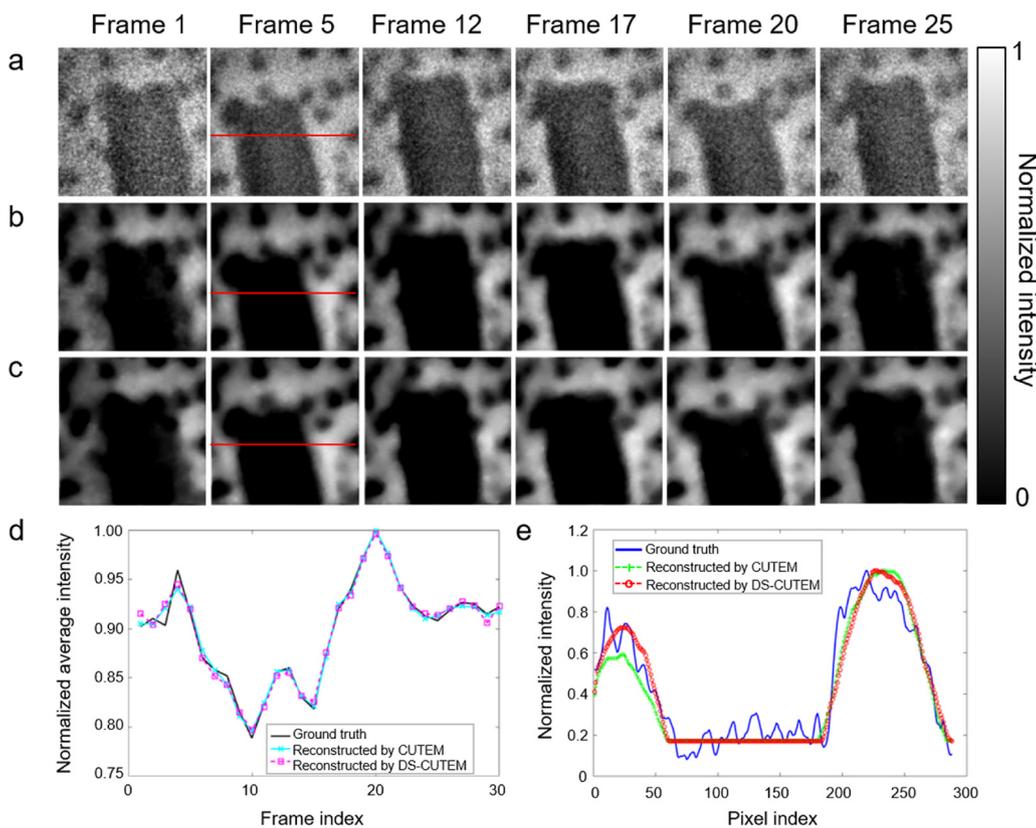


Fig. 6. Reconstruction of the cantilever scene using CUTEM and DS-CUTEM. (a), Representative frames of the cantilever scene as the ground truth [adapted from (Flannigan et al., 2009) with permission]. (b) and (c), Corresponding representative frames of reconstruction using CUTEM and DS-CUTEM, respectively. (d), Change in normalized average intensity across all frames of cantilever scene and the two reconstructed results. (e), Comparison of normalized intensity profiles of a horizontal line in the cantilever scene and the reconstruction results [marked by the red solid line in Frame 5 in (a)–(c)]. A frame-by-frame comparison of the two reconstructed datacubes with the ground truth is shown in Media 2.

into the spatiotemporal domain (Liang and Wang, 2018). By implementing the multiple-angle projections in the proposed configurations and by leveraging various CT reconstruction algorithms, the spatial and temporal resolutions of the CUTEEM and DS-CUTEEM systems could be further improved.

Acknowledgements

The authors thank Dr. Jingdan Liu for close reading of the manuscript. This work was supported in part by Natural Sciences and Engineering Research Council (NSERC) of Canada via programs of Discovery [RGPIN-2017-05959 (to J.L.) and RGPIN-2015-06133 (to A.Y.)], Discovery Accelerator Supplement [RGPAS-507845-2017 (to J.L.)], and Canada Research Chair (to A.Y.), and by Fonds de recherche du Québec–Nature et technologies (FRQNT) via the programs of Établissement de nouveaux chercheurs universitaires [2019-NC-252960 (to J.L.) and A.Y.] and Projets de recherche en équipe (to A.Y.). A patent application for the proposed methods has been filed.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.micron.2018.11.003>.

References

- Abolghasemi, V., Ferdowsi, S., Makkiabadi, B., Sanei, S., 2010. On optimization of the measurement matrix for compressive sensing. *Signal Processing Conference* 427–431.
- Abolghasemi, V., Ferdowsi, S., Sanei, S., 2012. A gradient-based alternating minimization approach for optimization of the measurement matrix in compressive sensing. *Signal Process.* 92, 999–1009.
- Baum, P., Yang, D.S., Zewail, A.H., 2007. 4D visualization of transitional structures in phase transformations by electron diffraction. *Science* 318, 788–792.
- Bernstein, D.S., 2005. *Matrix Mathematics: Theory, Facts, and Formulas With Application to Linear Systems Theory*. Princeton University Press.
- Bioucas-Dias, J.M., Figueiredo, M.A., 2007. A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Image Process.* 16, 2992–3004.
- Bostanjoglo, O., 2002. High-speed electron microscopy. *Adv. Imag. Elect. Phys.* 121, 1–52.
- Bostanjoglo, O., Weingärtner, M., 1997. Pulsed photoelectron microscope for imaging laser-induced nanosecond processes. *Rev. Sci. Instrum.* 68, 2456–2460.
- Bostanjoglo, O., Kornitzky, J., Tomow, R., 1991. High-speed electron microscopy of laser-induced vaporization of thin films. *J. Appl. Phys.* 69, 2581–2583.
- Bücker, K., Picher, M., Crégut, O., LaGrange, T., Reed, B., Park, S., Masiel, D., Banhart, F., 2016. Electron beam dynamics in an ultrafast transmission electron microscope with Wehnelt electrode. *Ultramicroscopy* 171, 8–18.
- Cavalleri, A., Dekorsy, T., Chong, H.H., Kieffer, J.-C., Schoenlein, R.W., 2004. Evidence for a structurally-driven insulator-to-metal transition in VO₂: A view from the ultrafast timescale. *Phys. Rev. B* 70, 161102.
- Elad, M., 2007. Optimized projections for compressed sensing. *IEEE Trans. Signal Process.* 55, 5695–5702.
- Elsayed-Ali, H., Norris, T., Pessot, M., Mourou, G., 1987. Time-resolved observation of electron-phonon relaxation in copper. *Phys. Rev. Lett.* 58, 1212.
- Feist, A., Bach, N., da Silva, N.R., Danz, T., Möller, M., Priebe, K.E., Domröse, T., Gatzmann, J.G., Rost, S., Schauss, J., 2017. Ultrafast transmission electron microscopy using a laser-driven field emitter: Femtosecond resolution with a high coherence electron beam. *Ultramicroscopy* 176, 63–73.
- Flannigan, D.J., Samartzis, P.C., Yurtsever, A., Zewail, A.H., 2009. Nanomechanical motions of cantilevers: direct imaging in real space and time with 4D electron microscopy. *Nano Lett.* 9, 875–881.
- Gao, L., Liang, J., Li, C., Wang, L.V., 2014. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature* 516, 74–77.
- Gedik, N., Yang, D.S., Logvenov, G., Bozovic, I., Zewail, A.H., 2007. Nonequilibrium phase transitions in cuprates observed by ultrafast electron crystallography. *Science* 316, 425–429.
- Gorkhover, T., Schorb, S., Coffee, R., Adolph, M., Foucar, L., Rupp, D., Aquila, A., Bozcek, J.D., Epp, S.W., Erk, B., 2016. Femtosecond and nanometre visualization of structural dynamics in superheated nanoparticles. *Nat. Photonics* 10, 93.
- Hunt, J., Driscoll, T., Mrozack, A., Lipworth, G., Reynolds, M., Brady, D., Smith, D.R., 2013. Metamaterial apertures for computational imaging. *Science* 339, 310–313.
- Kellman, P., McVeigh, E.R., 2005. Image reconstruction in SNR units: a general method for SNR measurement. *Magn. Reson. Med.* 54, 1439–1447.
- Kim, J.S., LaGrange, T., Reed, B.W., Taheri, M.L., Armstrong, M.R., King, W.E., Browning, N.D., Campbell, G.H., 2008. Imaging of transient structures using nanosecond in situ TEM. *Science* 321, 1472–1475.
- Koller, R., Schmid, L., Matsuda, N., Niederberger, T., Spinoulas, L., Cossairt, O., Schuster, G., Katsaggelos, A.K., 2015. High spatio-temporal resolution video with compressed sensing. *Opt. Express* 23, 15992–16007.
- LaGrange, T., Armstrong, M., Boyden, K., Brown, C., Campbell, G., Colvin, J., DeHope, W., Frank, A., Gibson, D., Hartemann, F., 2006. Single-shot dynamic transmission electron microscopy. *Appl. Phys. Lett.* 89, 044105.
- LaGrange, T., Reed, B.W., Santala, M.K., McKeown, J.T., Kulovits, A., Wiezorek, J.M., Nikolova, L., Rosei, F., Siwick, B.J., Campbell, G.H., 2012. Approaches for ultrafast imaging of transient materials processes in the transmission electron microscope. *Micron* 43, 1108–1120.
- Liang, J.Y., Wang, L.H., 2018. Single-shot ultrafast optical imaging. *Optica* 5, 1113–1127.
- Liang, J.Y., Ma, C., Zhu, L., Chen, Y., Gao, L., Wang, L.H., 2017. Single-shot real-time video recording of a photonic mach cone induced by a scattered light pulse. *Sci. Adv.* 3, e1601814.
- Liang, J.Y., Zhu, L.R., Wang, L.H., 2018. Single-shot real-time femtosecond imaging of temporal focusing. *Light-Sci. Appl.* 7, 42.
- Llull, P., Liao, X., Yuan, X., Yang, J., Kittle, D., Carin, L., Sapiro, G., Brady, D.J., 2013. Coded aperture compressive temporal imaging. *Opt. Express* 21, 10526–10545.
- Lustig, M., Donoho, D., Pauly, J.M., 2007. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* 58, 1182–1195.
- Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M., 2008. Compressed sensing MRI. *IEEE Signal Process. Mag.* 25, 72–82.
- Meurig, T.J., 2009. *4D Electron Microscopy: Imaging in Space and Time*. World Scientific.
- Musumeci, P., Cesar, D., Maxson, J., 2017. Double-shot MeV electron diffraction and microscopy. *Struct. Dyn.* 4, 044025.
- Reed, B.W., 2017. *Temporal Compressive Sensing Systems*. US. 20180136449 A1.
- Reed, B.W., Park, S.T., Bloom, R.S., Masiel, D.J., 2017. Compressively sensed video acquisition in transmission electron microscopy. *Microsc. Microanal.* 23, 84–85.
- Sjodin, T., Petek, H., Dai, H.-L., 1998. Ultrafast carrier dynamics in silicon: A two-color transient reflection grating study on a (111) surface. *Phys. Rev. Lett.* 81, 5664.
- Stevens, A., Kovarik, L., Abellan, P., Yuan, X., Carin, L., Browning, N.D., 2015. Applying compressive sensing to TEM video: a substantial frame rate increase on any camera. *Adv. Struct. Chem. Imaging* 1, 10.
- Strong, D., Chan, T., 2003. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Probl.* 19, S165.
- Tsiligianni, E.V., Kondi, L.P., Katsaggelos, A.K., 2014. Construction of incoherent unit norm tight frames with application to compressed sensing. *IEEE Trans. Inf. Theory* 60, 2319–2330.
- Vogel, C.R., Oman, M.E., 1996. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.* 17, 227–238.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612.
- Yurtsever, A., Baskin, J.S., Zewail, A.H., 2012. Entangled nanoparticles: Discovery by visualization in 4D electron microscopy. *Nano Lett.* 12, 5027–5032.
- Yurtsever, A., Zewail, A.H., 2012. Direct visualization of near-fields in nanoplasmonics and nanophotonics. *Nano Lett.* 12, 3334–3338.
- Zewail, A.H., 1990. The birth of molecules. *Sci. Am.* 263, 76–83.
- Zewail, A.H., 2000. Femtochemistry: atomic-scale dynamics of the chemical bond using ultrafast lasers (nobel lecture). *Angew. Chem. Int. Ed.* 39, 2586–2631.