

Space- and intensity-constrained reconstruction for compressed ultrafast photography

LIREN ZHU,[†] YUJIA CHEN,[†] JINYANG LIANG,[†] QIAOFENG XU, LIANG GAO, CHENG MA, AND LIHONG V. WANG*^{*}

Optical Imaging Laboratory, Department of Biomedical Engineering, Washington University in St. Louis, Campus Box 1097, One Brookings Drive, St. Louis, Missouri 63130, USA

*Corresponding author: lhwang@wustl.edu

Received 11 February 2016; revised 20 May 2016; accepted 21 May 2016 (Doc. ID 259326); published 30 June 2016

The single-shot compressed ultrafast photography (CUP) camera is the fastest receive-only camera in the world. In this Letter, we introduce an external CCD camera and a space- and intensity-constrained (SIC) reconstruction algorithm to improve the image quality of CUP. The external CCD camera takes a time-unsheared image of the dynamic scene. Unlike the previously used unconstrained algorithm, the proposed algorithm incorporates both spatial and intensity constraints based on the additional prior information provided by the external CCD camera. First, a spatial mask is extracted from the time-unsheared image to define the zone of action. Next, an intensity threshold is determined based on the similarity between the temporally projected image of the reconstructed datacube and the time-unsheared image. Both simulation and experimental studies show that the SIC reconstruction improves the spatial resolution, contrast, and general quality of the reconstructed image. © 2016 Optical Society of America

OCIS codes: (100.0118) Imaging ultrafast phenomena; (110.1758) Computational imaging; (110.3010) Image reconstruction techniques.

<http://dx.doi.org/10.1364/OPTICA.3.000694>

High-speed imaging technologies have been developed continuously for more than a century, going back to Talbot's recording of a spinning disk in the 1850s. The acquisition frame rates have been increased from ~50 frames per second with an intermittent camera [1] to one billion frames per second with a small number of frames, using a state-of-the-art electronic design with CCD and CMOS sensors [2]. However, with these electronic sensors, further increase in frame rate is impeded by the on-chip storage and electronic readout speed. The advent of the streak camera breaks this speed limit at the expense of the imaging dimensions—the use of a narrow entrance slit restrains its imaging to one spatial dimension.

To enable two-dimensional (2D) ultrafast imaging, different technologies have been developed. Pump-probe measurement is currently the predominant approach for capturing transient events [3–5]. Despite its widespread applications, this method requires that the transient events be precisely repeatable.

Single-shot ultrafast 2D imaging techniques have been developed [6–8]. For example, using burst illumination produced by a spatiotemporally modulated ultrashort laser pulse, the sequentially timed all-optical mapping photography camera has enabled trillion-frames-per-second photography with a sequence depth of up to six frames. However, its reliance on the specialized active illumination rules out imaging of chemiluminescent and photoluminescent objects. This challenge can be overcome by utilizing the streak camera as a 2D imager. Recent approaches include using a tilted lenslet array [9] or a 2D pinhole array [10] to achieve parallel streak framing. However, these methods suffer from significant throughput loss.

To overcome these limitations, we have developed compressed ultrafast photography (CUP), a new computational ultrafast imaging technology that can capture transient dynamic events at 100 billion frames per second in a single camera exposure with a sequence depth of hundreds of frames [11]. CUP synergistically combines two technologies: the streak camera and compressed sensing (CS). Unlike other streak-camera-based ultrafast imagers, CUP uses a fully opened entrance slit onto the streak camera. In addition, the dynamic scene is spatially encoded with a pseudorandom binary mask through a digital micromirror device (DMD). Given the spatiotemporal sparsity of the dynamic scene, which holds in many if not most natural scenes, a CS-based reconstruction algorithm can successfully decode the spatiotemporal mixing in the vertical axis of the streak camera and retrieve spatiotemporal information.

Using CUP, we have visualized many transient light-speed phenomena [11], including the propagation, reflection, and refraction of a short laser pulse in space, faster-than-light propagation of non-information, and color-resolved fluorescent excitation and emission. Recently, by leveraging the time-of-flight (ToF) of light signals backscattered from a three-dimensional object, CUP has also been used for dynamic volumetric imaging [12].

Currently, CUP relies on the unconstrained two-step iterative shrinkage/thresholding (TwIST) algorithm [13] to reconstruct the event datacube. The reconstructed image resolution is degraded by approximately a factor of 2 by the temporal shearing operation in the streak camera and reconstruction. In this Letter, we report the incorporation of additional prior information from an external CCD camera, which records a time-unsheared image of the dynamic scene. With the superior spatial resolution provided by the external CCD, we formulate a space- and

intensity-constrained (SIC) reconstruction algorithm to fully utilize this additional view of the scene, with the goals of improving image resolution, mitigating low-intensity artifacts, and boosting the general image quality of CUP.

In the conventional CUP system, the observed dynamic scene is spatially encoded by the DMD, temporally sheared by the streak tube, and then spatiotemporally integrated by the internal CCD camera inside the streak camera. As discussed previously [11], to reconstruct the original scene, one needs to solve the following inverse problem:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \{0.5 \|\mathcal{TSC}(\mathbf{x}) - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_{\text{TV}}\}, \quad (1)$$

where \mathbf{x} is the intensity distribution of the dynamic scene, \mathcal{T} is the spatiotemporal integration operator, \mathcal{S} is the temporal shearing operator, \mathcal{C} is the encoding operator that comes from the DMD, \mathbf{y} is the streak camera measurement, $\|\cdot\|_2$ denotes the ℓ^2 norm, $\|\cdot\|_{\text{TV}}$ denotes the total variation (TV) norm, and λ is the regularization parameter that tunes the ratio between the measurement fidelity term and the TV-based regularization term.

In our proposed method (outlined in Fig. 1), an external CCD camera records an unsheared spatiotemporally integrated image of the dynamic scene. Without the blurring caused by the temporal shearing in the streak tube, this additional perspective on the dynamic scene enjoys a better spatial resolution. To leverage this advantage, we first extract a spatial mask from the time-unsheared image, which is later used to define the zone of action in image reconstruction. A fairly standard grayscale segmentation approach is employed in this step: First, an adaptive local thresholding algorithm [14] is applied to the image with a 15-pixel circular local window to get an initial binary mask. Then, a 5-pixel median filter is applied to remove salt-and-pepper noise on the binary image. The resultant 2D binary mask is then used as a spatial constraint in the optimization framework so that pixels outside the mask are not updated during optimization and remain zero. Such a spatial constraint improves the spatial resolution of the reconstructed datacube and accelerates the reconstruction procedure by reducing the degrees of freedom of the underlying object function.

Second, to reduce the low-intensity artifacts in the reconstructed datacube, we introduce an intensity threshold constraint, based on the time-unsheared image, in the optimization algorithm. Taking advantage of the fact that an iterative shrinkage/thresholding optimizer can impose convex set constraints at each

iteration without losing its convergence properties [15], we apply an intensity threshold to the optimization problem in Eq. (1). Together with a spatial constraint, we have the new SIC solver modified from Eq. (1):

$$\mathbf{x}_s = \arg \min_{\mathbf{x} \in M, \mathbf{x} > s} \{0.5 \|\mathcal{TSC}(\mathbf{x}) - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_{\text{TV}}\}. \quad (2)$$

Here, M is the set of possible solutions confined by the spatial mask and s is the intensity threshold. Mathematically, the spatial mask forces background values to be zero and the intensity threshold ensures that all reconstructed pixel values are either greater than s or zero.

The optimal threshold is chosen based on the similarity between the external CCD camera image and the temporally integrated image of the reconstructed datacube. We use the root-mean-square error (RMSE) as the similarity criterion. If we denote \mathbf{y}_0 as the measured external CCD camera image and \mathbf{x}_s as the optimized solution with a threshold s , the optimal threshold is chosen by the following formula:

$$\hat{s} = \arg \min_s [1/N \cdot \|\mathcal{T}(\mathbf{x}_s) - \mathbf{y}_0\|_2^2]^{1/2}, \quad (3)$$

where the objective function is the RMSE and N is the total number of pixels in \mathbf{y}_0 . Since this optimization problem is based on a separate parameterized optimization problem (i.e., solving for \mathbf{x}_s with a given s), common gradient-based minimization methods become ineffective. We therefore employ a simple grid search method to find an approximation. More specifically, we first solve Eq. (2) with $s = 0$ and then find the largest pixel value in the solution. Values between 0 and 0.01 of this maximum pixel value are considered candidates for the optimal threshold. Eleven evenly distributed threshold values are then tried, and the RMSE criteria are calculated for each threshold. The result for the minimal RMSE is then chosen as the final reconstruction result.

Since multiple intermediate solutions of Eq. (2) are needed in our method, a fast algorithm becomes more desirable for SIC reconstruction. The fast iterative shrinkage/thresholding algorithm [15] has a reportedly faster convergence rate than the previously employed TwIST algorithm in solving general TV-regularized ℓ^2 -norm minimization problems, such as the one stated in Eq. (2), and hence is used in SIC reconstruction. To further accelerate our reconstruction method, we also implement the algorithm with the CUDA parallel programming framework on a single Tesla K40c graphic processing unit (GPU), reducing typical reconstruction time from tens of minutes to seconds.

We first validated our reconstruction method on numerically simulated data. A 200-by-200 Shepp-Logan (S-L) phantom was used as the base image. The simulated dynamic scene contained 10 frames, with the S-L phantom moving from left to right at four pixels per frame and flashing at the third, fifth, and eighth frames. The other frames are left black (set to zeros). The streak camera measurement was generated according to the forward model, and 1% Gaussian white noise was added. Similarly, the CCD measurement was generated by integrating the original dynamic scene along the time axis. To demonstrate the advantages of our method, the dynamic scene was reconstructed using both the conventional TwIST-based unconstrained reconstruction method and our proposed SIC reconstruction method.

Figure 2 shows the simulation results. The similarity metric varies when we tune the intensity threshold value, as shown in Fig. 2(a). In this particular case, a threshold of 0.008 gives us the best similarity and the datacube generated by this threshold

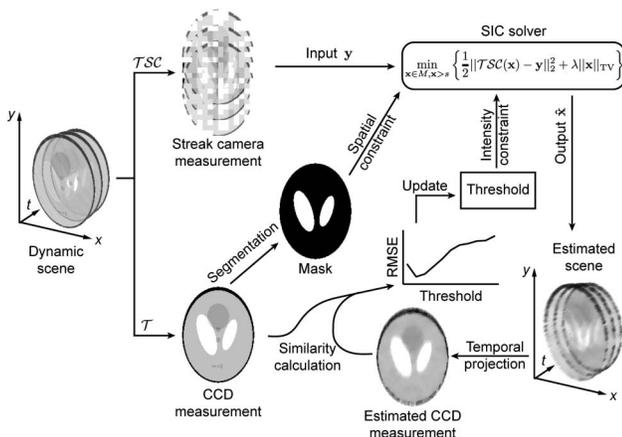


Fig. 1. Operation principle of the space- and intensity-constrained (SIC) reconstruction for compressed ultrafast photography (CUP).

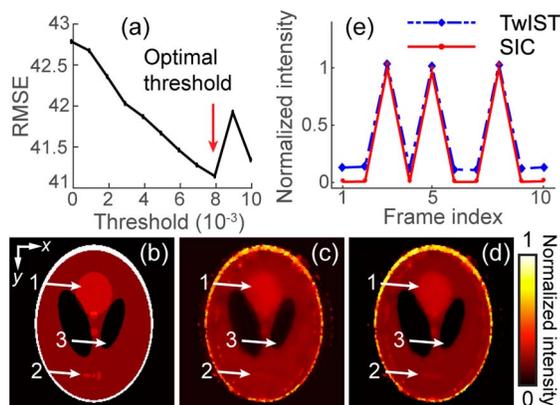


Fig. 2. Results of the numerical simulation. (a) Root-mean-square error (RMSE) as the similarity metric versus tested threshold values. (b) Fifth frame of the ground truth dynamic scene. (c) As (b), but shows the TwiST reconstructed result. (d) As (b) and (c), but shows the SIC reconstructed result. (e) Change in average normalized intensity across ten frames of the two reconstructed datacubes.

is chosen to be the final reconstruction result. Figures 2(b)–2(d) show the same (the fifth) frame of the ground truth, the TwiST reconstructed result, and the SIC reconstructed result. Focusing on Region 1, the large bright patch demonstrates better contrast in the image produced by the SIC reconstruction than that by the conventional TwiST reconstruction. In Region 2, the SIC method successfully recovers the small bright spot presented in the ground truth, while the conventional TwiST method fails to reconstruct this feature. The boundary between the dark and bright patches in Region 3 is also more prominent in our result than in the TwiST result. To compare the reconstruction quality of the two methods across frames, we plot the average normalized intensity against the frame index in Fig. 2(e). The SIC reconstruction leaves much less residual signal in the supposed black frames than the unconstrained TwiST method, demonstrating a better reconstruction performance in the time domain.

To experimentally validate our method, we upgraded the first-generation CUP system (Fig. 3). The dynamic scene is first imaged by a camera zoom lens (focal length 18–55 mm) to an intermediate image plane. Then a beam splitter divides the light into two directions. The reflected light is directly imaged by an external CCD camera. The transmitted light is passed to a DMD by a $4f$ imaging system with a tube lens and an objective lens. A pseudorandom binary pattern is programmed onto the DMD to spatially encode the dynamic scene. Collected by the same objective lens, the encoded scene is further imaged to the wide-open entrance slit of a streak camera. Inside the streak camera, the incident light is first imaged to a photocathode where light is converted into photoelectrons. After initial acceleration, these photoelectrons are sheared by a sweep voltage in the vertical axis, according to the ToF (inset in Fig. 3). Then the temporally sheared photoelectrons bombard a microchannel plate, where the current is amplified by generating secondary electrons. A phosphor screen converts the electrons back into light. An internal CCD camera then images the phosphor screen and compressively records the spatially encoded, temporally sheared dynamic scene in a single 2D image.

With this upgraded CUP system, we imaged a dynamic scene, namely a laser beam sweeping across a car-model target.

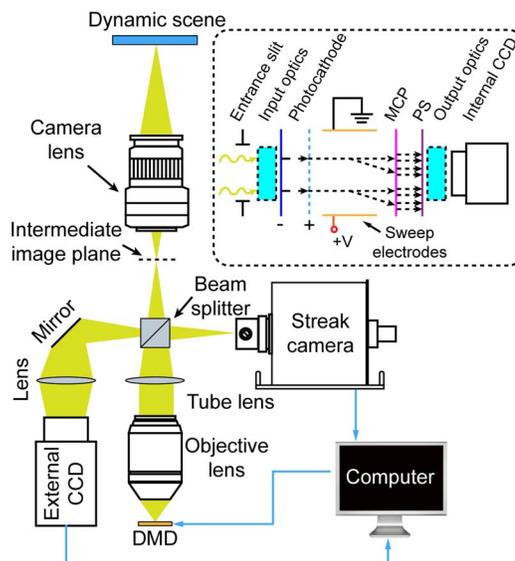


Fig. 3. System schematic. Inset: detailed illustration of the principle of the streak camera. DMD, digital micromirror device; MCP, microchannel plate; PS, phosphor screen; V, sweep voltage. Equipment details: camera lens, Nikon, $f = 18\text{--}55$ mm; DMD, Texas Instruments, LightCrafter 3000; external CCD camera, Point Gray, GS3-U3-28S4M-C; internal CCD, Hamamatsu, ORCA-R2; objective lens, Olympus, UPLSAPO $4\times$; streak camera, Hamamatsu, C7700; tube lens, Thorlabs, AC254-150-A.

A solid-state pulsed laser (532 nm wavelength, 7 ps pulse duration) was the light source. The laser beam was first passed through an engineered diffuser and illuminated the target at an oblique angle of $\sim 30^\circ$ with respect to the surface normal. The CUP system was placed perpendicular to the target's surface to collect the scattered photons. The system speed was 100 billion frames per second, achieved by setting the streak camera's shearing velocity to 1.32 mm/ns.

Similar to the simulation study, we reconstructed the dynamic scene using both the unconstrained TwiST and the new SIC reconstruction method. Figures 4(a) and 4(b) show temporally projected images of the reconstructed datacubes. A frame-by-frame comparison of the reconstructed datacubes from both methods is provided in Visualization 1. In the spatial domain, the results of the SIC method illustrate sharper boundaries. Figure 4(c) shows signal changes across time at the same group of pixels circled in Figs. 4(a) and 4(b). The full width at half-maximum (FWHM) of the temporal response is reduced from ~ 60 ps in the TwiST result to ~ 50 ps in the SIC result. Because the laser pulse width (7 ps) is sufficiently small, we can conclude that the SIC method improves temporal resolution by $\sim 17\%$. While the sharper boundaries are likely the outcome of the spatial constraint, the improved temporal resolution is largely due to the optimized intensity constraint. Figures 4(d) and 4(e) show the $x - y - t$ volumetric renderings of the datacubes reconstructed by the TwiST and SIC methods, respectively.

To further demonstrate the strength of our method, we imaged a picosecond laser pulse traveling in scattering air. The light source was the same laser used in the car model experiment. Figures 5(a) and 5(b) compare single frames extracted from the TwiST- and SIC-reconstructed datacubes. Visualization 2 is a frame-by-frame comparison of the reconstructed results from

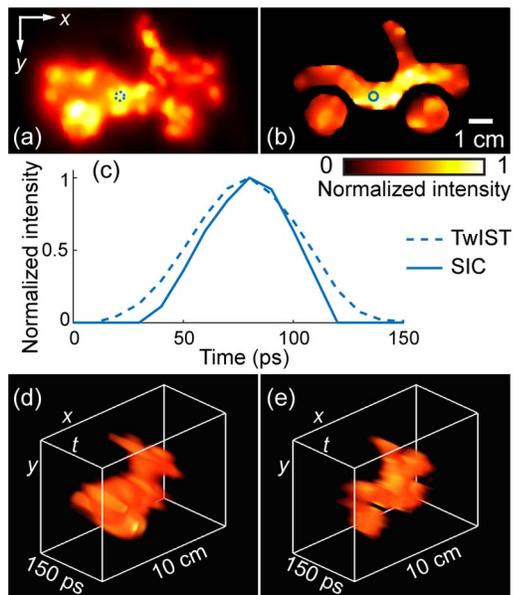


Fig. 4. Car model experimental results. (a) Temporally projected image of the TwIST-reconstructed datacube. (b) As (a), but showing the datacube from the SIC reconstruction. A frame-by-frame comparison of the two reconstructed datacubes is shown in Visualization 1. (c) Signal evolution at the circled pixels in (a) and (b). (d) An $x-y-t$ volumetric rendering of the TwIST-reconstructed datacube. (e) As (d), but showing the SIC-reconstructed datacube.

both methods. Improvement of image quality within the masked area (marked by the white curves) can be clearly observed in the SIC result. As shown in Fig. 5(c), the signal profile along the propagation direction is especially narrowed in the SIC result compared to that in the TwIST result, from a FWHM of 25.0 to 10.4 mm. This improvement resulted from a combination of improvements in the spatial and temporal resolutions.

Our previous ToF-CUP technique [12] uses a similar system setup, with one streak camera channel and one external CCD channel. However, ToF-CUP reconstruction only overlays the grayscale time-unsheared image on the TwIST-reconstructed datacube as a postprocessing method. The SIC method, on the contrary, extracts

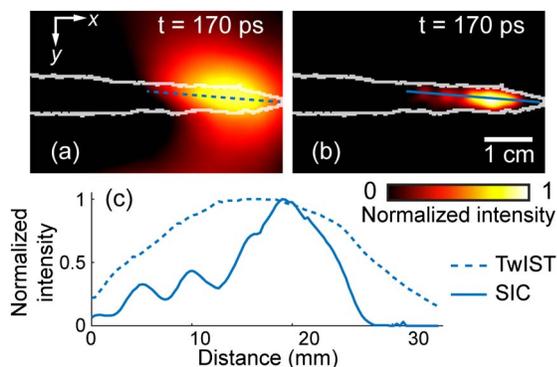


Fig. 5. Propagating picosecond laser pulse results. (a) TwIST-reconstructed frame at $t = 170$ ps. (b) SIC-reconstructed frame at $t = 170$ ps. White edges in both panels indicate the zone of action identified by the external CCD camera, which was not used in the TwIST reconstruction method. A frame-by-frame comparison of the two reconstructed datacubes is shown in Visualization 2. (c) Image profiles along the indicated lines labeled in (a) and (b).

a 2D mask and incorporates it as a spatial constraint in the optimizer. Moreover, by adding and optimizing an intensity constraint, SIC finds an adaptive spatiotemporal mask based on the recovered datacube. The combined effect of spatial and intensity constraints not only eliminated artifacts in the expected black regions, but also improved image quality inside the zone of action, as demonstrated in both simulated and experimental results. Naturally, by applying the SIC reconstruction method to the ToF-CUP system, one can expect even more accurate ToF information across the object surface, with better spatial resolution.

In conclusion, our SIC reconstruction method incorporates an external CCD camera that captures another perspective of the dynamic scene. Our new method exploits the resultant additional prior information by extracting spatial and intensity constraints from the external CCD image. As demonstrated with both numerical simulation and experimental data, the SIC reconstruction method recovers the dynamic scene with sharper boundaries, higher feature contrast, fewer low-intensity artifacts, and, in general, better image quality than the unconstrained TwIST reconstruction method in previous CUP technologies. Although it requires solving the original optimization problem multiple times, our SIC reconstruction leverages a faster reconstruction algorithm as well as current advances in computational hardware and GPU parallel computing technologies to reduce reconstruction time.

Funding. National Institutes of Health (NIH) (DP1 EB016986, R01 CA186567).

Acknowledgment. The authors would like to thank Prof. James Ballard for reading the manuscript.

[†]These authors contributed equally to this work.

REFERENCES

1. P. W. W. Fuller, *J. Imaging Sci.* **57**, 293 (2009).
2. Invisible Vision Ltd, [http://www.invisiblevision.com/pdf/UBSi_1Bn_fps_Camera\).pdf](http://www.invisiblevision.com/pdf/UBSi_1Bn_fps_Camera).pdf).
3. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, *Nat. Commun.* **3**, 745 (2012).
4. C. B. Schaffer, N. Nishimura, E. N. Glezer, A. M. T. Kim, and E. Mazur, *Opt. Express* **10**, 196 (2002).
5. G. Gariépy, N. Krstajić, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio, *Nat. Commun.* **6**, 6021 (2015).
6. Z. Li, R. Zgadzaj, X. Wang, Y.-Y. Chang, and M. C. Downer, *Nat. Commun.* **5**, 3085 (2014).
7. K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirose, H. Liao, T. Ushida, K. Goda, F. Kannari, and I. Sakuma, *Nat. Photonics* **8**, 695 (2014).
8. K. Goda, K. K. Tsia, and B. Jalali, *Nature* **458**, 1145 (2009).
9. B. Heshmat, G. Satat, C. Barsi, and R. Raskar, in *CLEO: Science and Innovations* (Optical Society of America, 2014), paper STu3E.7.
10. H. Shiraga, M. Nakasuji, M. Heya, and N. Miyanaga, *Rev. Sci. Instrum.* **70**, 620 (1999).
11. L. Gao, J. Liang, C. Li, and L. V. Wang, *Nature* **516**, 74 (2014).
12. J. Liang, L. Gao, P. Hai, C. Li, and L. V. Wang, *Sci. Rep.* **5**, 15504 (2015).
13. J. M. Bioucas-Dias and M. A. T. Figueiredo, *IEEE Trans. Image Process.* **16**, 2992 (2007).
14. N. Phansalkar, S. More, A. Sabale, and M. Joshi, in *International Conference on Communications and Signal Processing (ICCSPP)* (IEEE, 2011), p. 218.
15. A. Beck and M. Teboulle, *IEEE Trans. Image Process.* **18**, 2419 (2009).