Optics Letters

Temporal compressive edge imaging enabled by a lensless diffuser camera

Ze Zheng,[1,†] Baolei Liu,[1,†,*] Jiaqi Song,[1] Lei Ding,[2] Xiaolan Zhong,[1,6] Lingqian Chang,[3] Xiaojun Wu,[4] David McGloin,[5] and Fan Wang[1]

[1]School of Physics, Beihang University, Beijing 100191, China
[2]School of Biomedical Engineering, Faculty of Engineering & IT, University of Technology Sydney, New South Wales 2007, Australia
[3]Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China
[4]School of Electronic and Information Engineering, Beihang University, Beijing 100191, China
[5]School of Natural and Computing Science, University of Aberdeen, King's College, Aberdeen, AB24 3FX, UK
[6]zhongxl@buaa.edu.cn
[†]These authors contributed equally to this Letter.
*liubaolei@buaa.edu.cn

Received 21 December 2023; revised 14 March 2024; accepted 1 May 2024; posted 2 May 2024; published 27 May 2024

Lensless imagers based on diffusers or encoding masks enable high-dimensional imaging from a single-shot measurement and have been applied in various applications. However, to further extract image information such as edge detection, conventional post-processing filtering operations are needed after the reconstruction of the original object images in the diffuser imaging systems. Here, we present the concept of a temporal compressive edge detection method based on a lensless diffuser camera, which can directly recover a time sequence of edge images of a moving object from a single-shot measurement, without further post-processing steps. Our approach provides higher image quality during edge detection, compared with the "conventional post-processing method." We demonstrate the effectiveness of this approach by both numerical simulation and experiments. The proof-of-concept approach can be further developed with other image post-processing operations or versatile computer vision assignments toward task-oriented intelligent lensless imaging systems.

https://doi.org/10.1364/OL.515429

Lenses have played an essential role in optical imaging systems over the past few centuries. Traditional imaging systems designed to obtain high-dimensional images are often bulky and expensive, such as multi-shot and scanning methods for imaging multispectral or three-dimensional (3D) objects [1]. The recent rise of the concept of lensless "diffuser camera" is driving the development of miniaturized and low-cost cameras [2]. By encoding the high-dimensional information of objects, the diffuser-assisted lensless imaging systems are able to recover multiple-dimensional images such as 3D depth imaging [3], multispectral/hyperspectral imaging [4–6], 3D fluorescence imaging [7], compressive temporal imaging [8,9], full-Stokes polarization imaging [10], and multi-modality edge enhancement imaging [11]. Generally, the starting point for such devices is the characterization of the diffuser, such as calibrating the point spread functions (PSFs) of a ground glass diffuser or phase mask. Then, a two-dimensional (2D) image is captured by putting the diffuser, instead of a lens, between an object and a 2D sensor. Finally, inverse algorithms are implemented to recover high-dimensional images. Recent advances in diffuser cameras have shown great promise in both applications and adding new imaging functionality. For example, in vivo lensless 3D microscopy has been achieved by using a specially designed phase mask [12], and multidimensional imaging, including spatial, spectral, and polarization dimensions, can be encoded by a liquid crystal metasurface diffuser [13]. In addition, programmable diffusers, such as reconfigurable particle assembly masks [14], show potential for developing tunable diffuser cameras.

After the images of an object are obtained, image post-processing is critical for acquiring additional useful information [15]. For example, edge detection has been widely used in advanced driving assistance systems and geographic environment monitoring [16,17]. Since the results of the diffuser camera are reconstructed from the raw data, a fusion of the traditional inverse algorithms used in diffuser cameras and image post-processing algorithms could generate the post-processed images directly from the raw data, without recovering the original images of the objects. Such an implicit image processing method, which is involved in computational optical imaging systems, would offer advantages over the post-processing method, offering higher signal-to-noise ratios (SNRs) of edge detection in the case of computational ghost imaging [18–20].

In this work, we propose a temporal compressive edge detection method based on a lensless diffuser camera (DiffuserEdgeCam), which can directly recover a sequence of edge images of an object at different time points from a single-shot measurement, without further post-processing steps. By only modifying the forward model matrix in the reconstruction, rather than modifying any steps in the experiment, we can incorporate the image post-processing steps into the reconstruction algorithm.

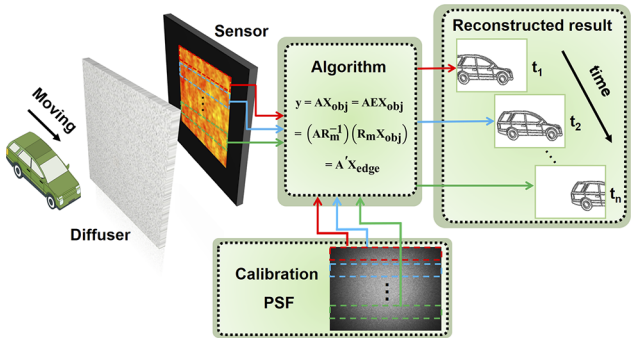

Fig. 1. Principle of the lensless diffuser camera for temporal compressive edge detection (i.e., Diffuser-eCam). The system encodes the temporal moving object in a 2D image on the chip CMOS sensor. The sensor works in a rolling shutter mode in which different rows of the CMOS sensor begin to be exposed at different times. For calibration, a 100 μm pinhole is placed at the object plane for the capture of the system's PSF, as shown at the bottom. By only modifying the forward model matrix A in the reconstruction, the edge images of the object can be directly reconstructed by inverse algorithms. The multi-frame edge images of the moving object are reconstructed from different rows of the measurement and corresponding rows of the PSF image.

To demonstrate this, we introduce an edge detection filter into the model matrix, which serves as a spatial filtering operation, allowing the inverse algorithms to directly reconstruct the edge images, without recovering their original images. With both numeral simulations and experimental results, we show that this method can improve the peak signal-to-noise ratios (PSNRs) and information entropy (IE) of the resulting images more than the "conventional post-processing method" using the same filter. Furthermore, we show how to recover multi-frame edge images of a moving object from a single-frame 2D measurement, by exploring the rolling shutter mode of the 2D sensor [8,9,21,22]. Other image post-processing operations can be studied in diffuser cameras or other computational imaging systems, to bypass the post-processing steps and enhance their resulting image quality. A diffuser camera is considered as a lensless imaging system that uses a pseudo-random diffuser instead of the traditional lenses to modulate the light field, as shown in Fig. 1. The sensor no longer receives an image of the object, but a blurry speckle pattern. In this work, we assume that the object is located within the angular memory effect region of the diffuser, in which the PSFs corresponding to different point sources do not change their shape and only translate in the image plane [23]. Another assumption is all the point sources within the object are incoherent with each other. Then, the 2D measurement I can be expressed as the convolution of the object O and the system's PSF P:

$$I(x,y) = O(x,y) * P(x,y), \tag{1}$$

where $*$ represents the convolution operation. For convenience, we can also express it in the form of matrix multiplication:

$$y = AX_{obj}, \tag{2}$$

where y is the column vector composed of intensity values of different pixels of I, X_{obj} is the column vector composed of each intensity value of O, A is the forward model matrix or calibration matrix that is related to the light modulation of the diffuser, and each of its columns can be obtained by translating and reshaping the diffuser's PSF. By using inverse algorithms, the original image X_{obj} can be reconstructed.

To extract the edges of the object, conventional methods require post-processing operations on the original image X_{obj}, which needs to be reconstructed in advance. Here we choose an edge detection operator R [19] to demonstrate this effect:

$$R = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3}$$

The desired result X_{edge} can be yielded by convolving X_{obj} and the edge detection operator R. It can also be expressed as the form of matrix multiplication:

$$X_{edge} = R * O(x,y) = R_m X_{obj}, \tag{4}$$

where R_m is the corresponding matrix of the operator R. We can use the identity matrix $E = R_m R_m^{-1}$ to modify Eq. (2):

$$y = AX_{obj} = AEX_{obj} = (AR_m^{-1})(R_m X_{obj}) = A' X_{edge}, \tag{5}$$

where $A' = AR_m{}^{-1}$ is the modified forward model matrix that can be generated before the reconstruction starts. Then, the compressed sensing algorithm, such as compressive sensing with total variation regularization [24] and the TwIST algorithm [25], can directly extract the object's edge X_{edge}:

$$\hat{X}_{edge} = arg \min_{X_{edge} \geq 0} \|y - A' X_{edge}\|_2^2 - \tau \|\Psi X_{edge}\|_1, \tag{6}$$

where Ψ is the linear change matrix that maps X_{edge} to the domain with sparse representation and τ is a regularization parameter that tunes the sparsity of the scene.

To reconstruct the images of a temporal moving object, we can further explore the rolling shutter mode of the CMOS sensor. The 2D resulting intensity distribution I can be expressed as the convolution of the PSF P and the object $O(x, y, t)$:

$$I(x,y) = O(x,y,t) * P(x,y). \tag{7}$$

The measurement y in the rolling shutter mode can be expressed as the form of temporal integration:

$$y(r,c) = \int_0^{r_{max}\Delta + T_e} S(t|r,c) I(r,c,t)\, dt, \tag{8}$$

where r and c are the row index and column index of the CMOS sensor, r_{max} is the max row index (r, c, and r_{max} are required to be positive integers), T_e is the exposure time, and Δt is the offset delay time of the sensor's rolling shutter mode. S is the temporal encoded function [9]:

$$S(t|r,c) = \begin{cases} 1, & \text{when} \lceil \frac{t - T_e}{\Delta t} \rceil \leq r \leq \lceil \frac{t}{\Delta t} \rceil + 1 \\ 0, & \text{else} \end{cases}. \tag{9}$$

Thus, the information of the temporal object is encoded into different rows of the single-shot raw data. By inputting the inverse algorithms with these different rows, we can directly reconstruct the multi-frame edge images of a moving object that corresponds to different time points.

To demonstrate our proposed method, we conduct the numerical simulation with static objects shown in Fig. 2. The original images, which are the gray scale versions of "hand" and "tumor

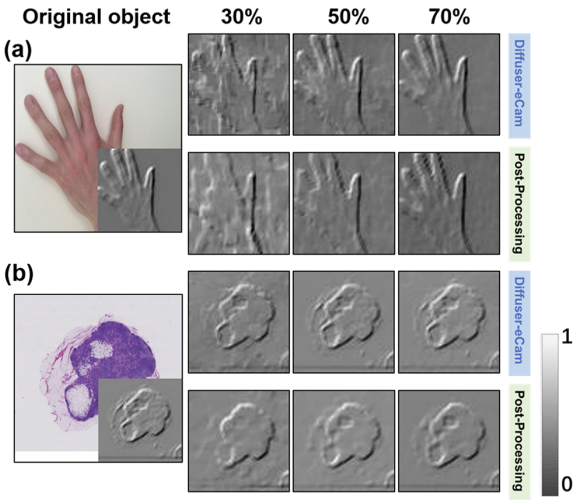

Fig. 2. Simulated results of edge detection by using the Diffuser-eCam and "conventional post-processing method" with sampling rates of 30%, 50%, and 70%. The original objects are the images of (a) "hand" and (b) "tumor tissue." The ground truth images are shown in the lower right corner of the corresponding original objects' images.

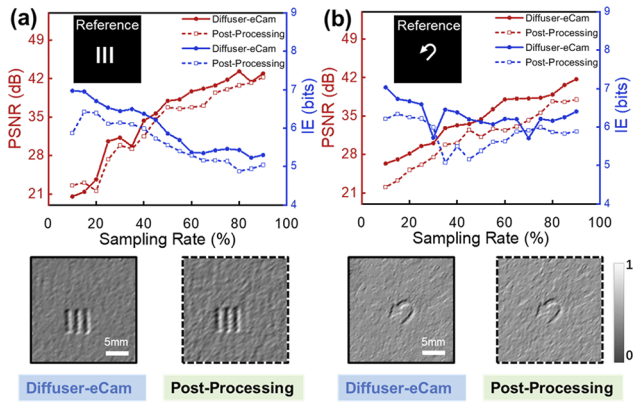

Fig. 3. Experimental results and comparison of the Diffuser-eCam and "conventional post-processing method" with the two objects: (a) three stripes and (b) the U-turn arrow. The example resulting images of two different objects are shown at the bottom of each subfigure, with a sampling rate of 50%. The reference images are shown in the upper left corner of each subfigure. The image quality of the results, at different sampling rates, is characterized by using PSNR and IE. The red lines represent the PSNRs, and the blue lines represent the IE values. The solid and dashed lines are the calculated values from the Diffuser-eCam and "conventional post-processing method," respectively.

tissue" from MATLAB's image data, are shown in the left column of Fig. 2. The ground truth edge images are embedded in the lower right corner of the corresponding original images. The results from the Diffuser-eCam and "conventional post-processing method" at different sampling rates are shown in Fig. 2 and indicated with Diffuser-eCam and "post-processing," respectively (see Supplement 1). The edge images of the Diffuser-eCam are directly reconstructed from Eq. (6). The "post-processing" method requires the reconstruction of the image of the objects first using the compressive sensing algorithm with the same prior according to Eq. (2) and then the convolution of the object images and the edge detection operator according to Eq. (4).

The image quality of the results, for both methods, increases with the sampling rate. It is evident that the results of the Diffuser-eCam have higher contrast and more details than those from the "post-processing method." For example, the edge detection results of "hand" and "tumor tissue" from the Diffuser-eCam are more prominent and have higher image contrast than the post-processing method. The reason for the phenomenon can be explained that objects' edges have much higher sparsity than the original objects' images [20]. It is also noted that the quality of edge detection, for the majority of real objects, is improved if the object's edge can be reconstructed directly, compared with the traditional procedure that needs to reconstruct the original object image in advance [18–20].

Then we conducted experiments to verify the advantages of the proposed method by using static transmission objects, as shown in Fig. 3. The collimated incoherent light first illuminated the objects. The transmitted light was modulated by a diffuser and then captured by a CMOS sensor (daA2500-14um, Basler), which is placed 1 cm behind the diffuser. The PSF is calibrated once as the objects are within the range of the optical angular memory effect (see Supplement 1 for more details).

Figure 3 compares the post-processing method and Diffuser-eCam for two different objects. The resulting images are evaluated by the PSNR and IE [26]. With a sampling rate of 50%, the results from the Diffuser-eCam have smoother and less noisy backgrounds than the post-processing method, as shown at the bottom of each subfigure. The PSNRs of the images from the two methods, at different sampling rates, are shown in red solid and dashed lines, respectively. Generally, the PSNRs for different objects increase as the sampling rate increases. Compared with the post-processing method (red dashed lines), the resulting images of the Diffuser-eCam have higher PSNRs and less noise (red solid lines). The corresponding values of IE are shown in the blue solid and dashed lines.

As shown in Fig. 3, the solid lines, whether red or blue, are almost above the dashed lines. In other words, the PSNRs and IE values of the results from the Diffuser-eCam are mostly higher than the "conventional post-processing" results. The edge is usually the place where the pixel value changes most obviously in gray scale. Better edge quality results in higher IE value and higher PSNR value that indicate lower noise. It implies that the contribution to the growth of the PSNR and the IE is mostly generated from the sharper edge rather than the more complex background. Thus, it can be concluded that the Diffuser-eCam has improved the image quality of the objects' edge, compared with the "conventional post-processing method."

One advantage of diffuser cameras is their ability to achieve compressive imaging that encodes a temporal scene into a single-shot measurement by using the rolling shutter mode of the sensor, in which the pixels in each row of the sensor are sequentially exposed at different times. To further demonstrate the reconstructed multi-frame edge images of a moving object from a single-shot 2D measured speckle, as shown in Fig. 4, we make use of a 3D-printed "car" as the object, as shown in the inset image of Fig. 4(a). The "car" is mounted on a translation stage (GCM-083904M, DHC) and is moved parallel to the diffuser camera. Figure 4(b) shows the directly captured image of the moving "car" by using the traditional lensed system with the

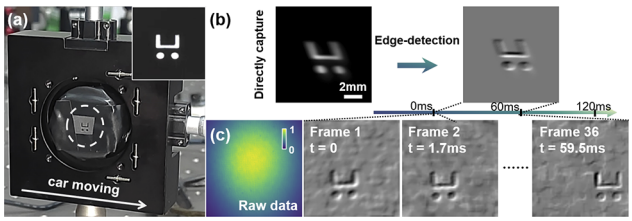

Fig. 4. Demonstration of the Diffuser-eCam for reconstructing multi-frame edge images of a moving object, from a single-shot 2D raw data. (a) 3D-printed "car" object is mounted on a translation stage and is moving laterally in the experiment. The "car" is about 3 mm × 3 mm, and its ground truth image is shown in the upper right corner of (a). The directly captured photo of the moving "car" by using a traditional lensed system, and the corresponding edge image is shown in (b). (c) 2D raw data captured by the Diffuser-eCam is shown on the left. The example reconstructed edge images (with a resolution of 65 × 65 pixels) corresponding to different time points (t = 0, 1.7 ms,. . . , 59.5 ms) are shown in the middle and right, which are selected from the 36 edge frames at 588 fps. Both the traditional lensed system and Diffuser-eCam use the same CMOS sensor (daA2500-14um, Basler), which is operated in the rolling shutter mode and has a maximum frame rate of 14 fps at the full resolution.

same CMOS sensor in the rolling shutter mode, and its post-processed edge image is shown on the right. The image of the "moving car" in Fig. 4(b) looks skewed, due to the rolling shutter effect while shooting a dynamic scene. Figure 4(c) shows the results of the Diffuser-eCam. The raw data is shown on the left, which was also captured in the rolling shutter mode. Since different rows of the measurement correspond to different time points, the resulting multi-frame edge images at certain time points can be reconstructed directly from the corresponding rows of the raw data by the Diffuser-eCam. We can clearly observe that the "car" is moving from the left to the right from the sequential 36 frames (see Supplement 1). The selected edge images of the moving "car" are shown in the middle and right of Fig. 4(c), corresponding to different time points (t = 0, 1.7 ms,. . . , 59.5 ms). Different from Fig. 4(b) and its edge image, the reconstructed results by the Diffuser-eCam show the clear and undistorted edge images at 588 frame-per-second (fps), which is enough to reveal that the moving "car" has an average velocity of 67.2 mm/s during the exposure (see Supplement 1). Here the Diffuser-eCam is able to directly reconstruct 36 frames edge images of the moving object at 588 fps, in which the employed CMOS sensor is operated at a much slower frame rate.

In conclusion, we have proposed and demonstrated temporal compressive edge imaging based on a lensless diffuser camera from a single-shot measurement. The diffuser camera encodes a temporal scene into a 2D image on the sensor. By only modifying the forward model matrix in inverse algorithms, the edge images of an object corresponding to different times can be directly reconstructed from different rows of the 2D raw measurement. The proposed method can achieve higher image quality, compared with the "conventional post-processing method" that convolves the retrieved object image with an edge detection operator. Thus, the Diffuser-eCam shows not only a new dimension of diffuser cameras for edge detection but also higher resulting image quality than the "conventional post-processing method."

One major advantage of the proposed method is that it does not require any change to the experimental setup or multiple measurements, but only modifying the forward model matrix during the reconstruction process. Therefore, our method would inspire further developments of diffuser cameras with the realization of other digital image processing tasks, such as high-/low-pass filtering, deconvolution, and denoising. Further complex vision tasks, such as face detection and fingerprint identification, can be also considered. The simple hardware system also makes the diffuser camera easily add new imaging dimensions, such as wavelength and polarization, into the imaging systems. Our scheme provides a new way to realize edge detection based on a lensless diffuser camera. We, therefore, anticipate that this work will open opportunities for developing smart lensless imaging systems with versatile vision tasks.

Funding. National Natural Science Foundation of China (U23A20481, 11804018, 62075004, 62275010); China Postdoctoral Science Foundation (2022M720347); Beijing Municipal Natural Science Foundation (1232027, 4212051); International Postdoctoral Exchange Fellowship Program (YJ20220241); Fundamental Research Funds for the Central Universities.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

REFERENCES

1. L. Gao and L. V. Wang, Phys. Rep. **616**, 1 (2016).
2. V. Boominathan, J. T. Robinson, L. Waller, *et al.*, Optica **9**, 1 (2022).
3. N. Antipa, G. Kuo, R. Heckel, *et al.*, Optica **5**, 1 (2018).
4. S. K. Sahoo, D. Tang, and C. Dang, Optica **4**, 1209 (2017).
5. X. Li, J. A. Greenberg, and M. E. Gehm, Optica **6**, 864 (2019).
6. K. Monakhova, K. Yanny, N. Aggarwal, *et al.*, Optica **7**, 1298 (2020).
7. K. Yanny, N. Antipa, W. Liberti, *et al.*, Light: Sci. Appl. **9**, 171 (2020).
8. G. Weinberg and O. Katz, Opt. Express **28**, 30616 (2020).
9. N. Antipa, P. Oare, E. Bostan, *et al.*, in *2019 IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2019), pp. 1–8.
10. N. Baek, Y. Lee, T. Kim, *et al.*, APL Photonics **7**, 116107 (2022).
11. L. Li, J. Ma, D. Sun, *et al.*, Opt. Express **31**, 22519 (2023).
12. J. K. Adams, D. Yan, J. Wu, *et al.*, Nat. Biomed. Eng. **6**, 617 (2022).
13. Y. Lei, Q. Zhang, Y. Guo, *et al.*, Photonics Res. **11**, B111 (2023).
14. J. R. Miller, C.-Y. Wang, C. D. Keating, *et al.*, ACS Nano **14**, 13038 (2020).
15. R. N. Strickland and M. Y. Aly, Opt. Eng. **24**, 683 (1985).
16. J. Canny, IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8**, 679 (1986).
17. X. Hu and Y. Wang, Catena **209**, 105840 (2022).
18. Z. Ye, P. Zheng, W. Hou, *et al.*, Opt. Lasers Eng. **159**, 107191 (2022).
19. H. Penketh, W. L. Barnes, and J. Bertolotti, Opt. Express **30**, 7035 (2022).
20. X. Liu, X. Yao, R. Lan, *et al.*, Opt. Express **23**, 33802 (2015).
21. M. Sheinin, Y. Y. Schechner, and K. N. Kutulakos, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 6437–6446.
22. M. Sheinin, D. Chan, M. O'Toole, *et al.*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 16324–16333.
23. I. Freund, M. Rosenbluh, and S. Feng, Phys. Rev. Lett. **61**, 2328 (1988).
24. L. Bian, J. Suo, Q. Dai, *et al.*, J. Opt. Soc. Am. A **35**, 78 (2018).
25. J. M. Bioucas-Dias and M. A. T. Figueiredo, IEEE Trans. Image Process **16**, 2992 (2007).
26. R. C. Gonzalez and R. E. Woods, in *Digital Image Processing* (Prentice-Hall Inc., 2006).