Single-shot wide-field imaging in reflection by using a single multimode fiber

Yifan Liu, Panpan Yu, Yijing Wu, et al.
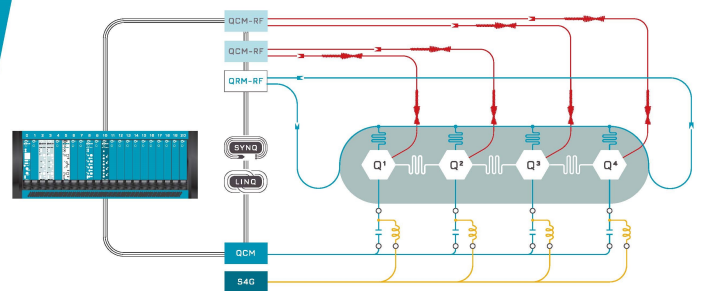
View Online   Export Citation   CrossMark

QBLOX

Integrates all
Instrumentation + Software
for Control and Readout of
Superconducting Qubits

visit our website >

AIP Publishing

Single-shot wide-field imaging in reflection by using a single multimode fiber

View Online   Export Citation   CrossMark

Yifan Liu,[1] Panpan Yu,[1] Yijing Wu,[1] Ziqiang Wang,[1] Yinmei Li,[1] Jinyang Liang,[2] Puxiang Lai,[3] and Lei Gong[1,a]

AFFILIATIONS

[1]Department of Optics and Optical Engineering, University of Science and Technology of China, Hefei 230026, China
[2]Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, Université du Québec, 1650 Boulevard Lionel-Boulet, Varennes, Québec J3X1P7, Canada
[3]Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong, China

[a]Author to whom correspondence should be addressed: leigong@ustc.edu.cn

ABSTRACT

A single multimode fiber (MMF) provides almost an ideal optical channel to constitute a hair-thin endoscope for minimally invasive biomedical imaging at depths in tissue, especially if the imaging operation can be performed with one single shot in reflection mode, which, however, remains challenging to date. In this work, we present single-shot wide-field reflectance imaging by using a single MMF as the illumination unit and imaging probe simultaneously. To achieve single-shot image capture, a reflection matrix of the fiber was built by a learning-assisted approach for the universal inverse conversion from the output amplitudes to the input amplitudes. The performance was tested by imaging more than 30 000 natural scenes projected by a digital micromirror device, and an averaged Pearson correlation coefficient over 0.84 with respect to the ground truth was achieved in the experiment. Furthermore, the ability to image dynamic scenes at a high frame rate of up to 180 frames per second was demonstrated together with real-time observation of a freely moving microneedle located at the distal end of the MMF. The proposed reflection-mode single-fiber imaging scheme paves the way for practical video-rate microendoscopy at depths in tissue in a minimally invasive manner.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0132123

Endoscopes provide a powerful and indispensable tool for medical diagnosis and industrial inspection in difficult-to-reach places. Benefitting from the high mechanical flexibility and small footprint, fiber-optic imaging systems are emerging as miniaturized endoscopes for minimally invasive biomedical applications, such as cellular probing[1] and *in vivo* brain imaging.[2–4] Most fiber imaging systems employ fiber bundles or multicore fibers as probes,[5–8] where each core acts as an individual pixel to convey image information. Nevertheless, multicore fiber endoscopes suffer from reoccurring image replicas due to the periodicity of multiple cores, while fiber bundles are limited in resolution by pixelation artifacts. In contrast, single multimode fiber (MMF) based endoscopes can offer a dramatic footprint reduction and a significant enhancement in spatial resolution,[2] achieving less invasion and higher quality images in endoscopy. Most such endoscopes actually consist of two fibers: a single-mode fiber for illumination and an MMF for image transmission.[9–12] This structure requires special coupling or fabrication. For endoscopes using only a single MMF, nearly all the systems have focused on transmission

geometry.[13–22] They require the imaged objects to be inserted between the illumination source and the fiber probe, which inevitably limits the applicable spaces in practical scenarios. If adopting reflection geometry, the imaging system must use the MMF to send illumination light and collect signal light simultaneously. In this case, current systems usually adopt a bucket detector to detect the signal light, and complex hardware is required for scanning imaging.[11,12,23–30] An endoscope equipped with a pixelated sensor offers wide-field imaging possibilities,[31,32] but the distortion of both illumination and signal light obstructs implementation. Overall, a reflection-mode single MMF endoscope is highly desirable for practical wide-field imaging; however, it still remains challenging to realize a simple and compact setup.

Light distortion due to mode dispersion within the MMF scrambles the transmitted image and yields seemingly random speckles at the output facet.[14,33,34] Handling the distortion becomes a key issue for single-fiber imaging, which can be achieved by wavefront shaping techniques. For instance, many studies have proven that the prior calibration of the MMF in the form of a transmission matrix (TM)

enables the recovery of object images.[11,12,18,23,24,31,32,34] Specifically, an image can be captured by sequential scanning of each pixel with the focused spot created by wavefront shaping using the TM. To detect the full complex field transmitted through the fiber, the TM calibration involves interferometric optics and is very susceptible to fiber bending. Apart from raster-scan imaging, one can also perform speckle imaging to see through an MMF with the calibrated TM. For instance, Choi *et al.* presented wide-field speckle imaging by a turbid lens imaging algorithm.[31] Lee *et al.* proposed synthetic confocal gated imaging through MMF.[32] These speckle imaging methods rely on detecting multiple frame fields to reconstruct an object image, preventing single-shot imaging and thus limiting the frame rate. In recent years, deep learning methods have started to be applied to single MMF imaging.[15–17,25,35,36] Distinguished from the TM-based methods, artificial neural networks are able to learn the nonlinear input–output relationship of the MMF with intensity-only detection, which is much simpler to implement. Moreover, deep learning methods could help improve the system's robustness to the environment. Nevertheless, convolutional artificial neural networks are typically limited to image scenes that belong to the same class as the training dataset. To address this limitation, an approach that directly learns the inverse TM was proposed for recovering natural scene images.[20] Despite these advances in learning-based methods, most works hitherto have focused on the transmission of images through the MMF. Thus far, single-shot wide-field imaging in reflection by using a single fiber has not yet been realized.

In this work, we demonstrate single-shot reflectance imaging by using a reflection-mode single-fiber system where the light source and the camera are located on the same side of the fiber and no lens or scanner is attached to the tip. Furthermore, a learning-based approach constructs a reflection matrix of the MMF for the inverse conversion from the output amplitude speckles to the input amplitude patterns under distorted illumination. After that, we are capable of recovering an object image from a single camera frame of the backscattered light with the single-fiber system. Distinguished from the existing reflection speckle imaging methods,[31,32] our single-shot imaging scheme enables endoscopic imaging at higher frame rates without measuring complex field information. By using a digital micromirror device (DMD) to project the target scenes, we demonstrate single-shot imaging of arbitrary natural scenes with an averaged Pearson correlation coefficient (PCC) > 0.84 and particularly imaging dynamic scenes at a frame rate of up to 180 frames per second (fps). In addition, we show that the proposed technique is capable of imaging with an enlarged field of view (FOV) with the same setup, which is realized by translating the fiber end to take images at different parts of the target scene. Finally, to demonstrate practical imaging, a microneedle is loaded at the distal end of the fiber and moved rapidly. The proposed technique achieves real-time observation of the moving object. The reflection-mode single-fiber imaging scheme paves the way for video-rate endoscopy in a minimally invasive manner, with potential benefits to practical biomedical applications.

Figure 1(a) illustrates the imaging geometry of the reflection-mode single-fiber endoscopic system. Our endoscope probe is simply a single MMF [1 m, numerical aperture (NA): 0.22, core diameter: 600 μm] with no lens or scanner attached to the tip. Coherent light from a laser source propagates through the MMF to illuminate the object being imaged at the distal end. The backscattered light is then
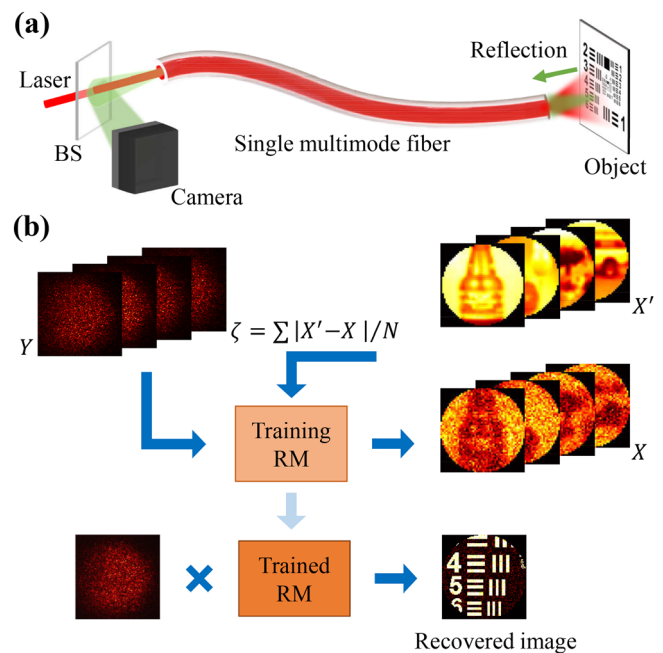

FIG. 1. Schematic overview of reflection-mode single-shot endoscopic imaging using a single MMF. (a) Imaging geometry of the single-fiber endoscopic system. (b) Process of reflection matrix training and image reconstruction from a single-shot backscattered speckle pattern.

collected by the same MMF and delivered to the proximal end, where a beam splitter separates the input and reflected light. A camera records the intensity of the backscattered light for image recovery. More details of the experimental setup can be found in the supplementary material. In fact, the optical scattering process produces a result where the MMF transforms the object information at the distal end into optical speckles at the proximal fiber end. Hence, the key to image recovery lies in the inverse mapping between the output amplitude speckles and the input amplitude of the object. Referring to the concept of a reflection matrix, the inverse problem can be described by

$$X = |R \cdot Y|, \tag{1}$$

where X is the amplitude of the imaged object, and Y is the amplitude of the output speckle field. It is noted that the reflection matrix R actually maps the amplitude-to-amplitude inversion, which can be taken as an approximate model of the inverse TM that correlates the input–output fields. In this case, the inverse mapping is nonlinear due to the intensity-only detection, without measuring the phase information at the output end of the fiber. Even so, it applies the TM theory that assumes that light is a monochromatic wave. If one wants to project two objects with identical amplitude functions but different phase functions, our method cannot reconstruct the correct images because this requires mapping the field-to-field inversion (i.e., conventional TM).

Here, we employ a learning-based approach to train the reflection matrix of the fiber under the illumination of distorted light, which is referred to as the TM optimization method first reported by Caramazza *et al.*[20] The training process is schematically shown in
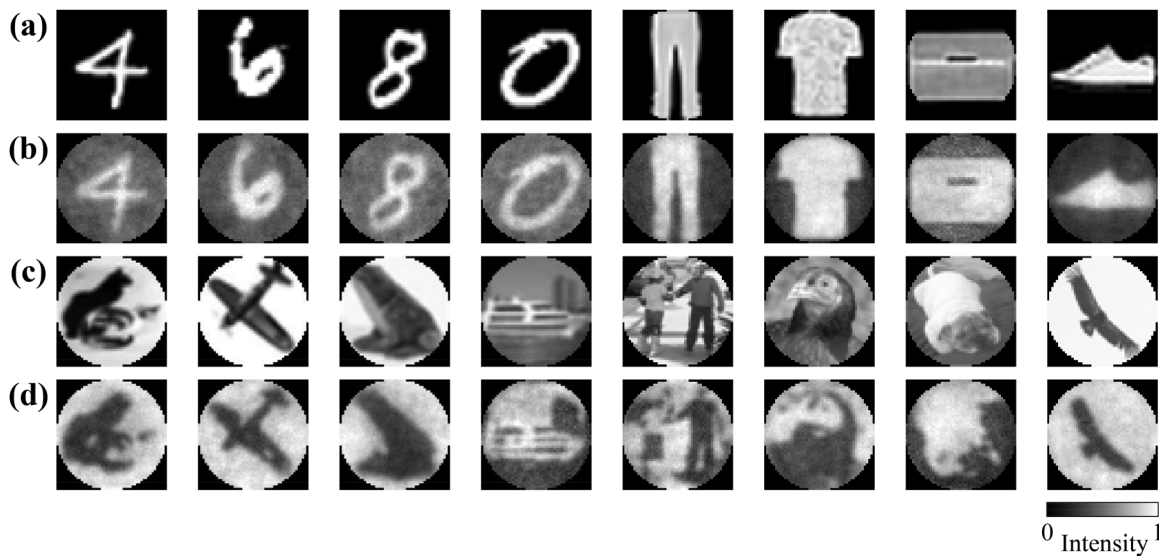
FIG. 2. Imaging testing with complex natural scenes. (a) and (c) The ground truth of the original images displayed on the DMD as the target scenes being imaged. The pictures come from various image datasets, including MNIST, Fashion-MNIST, CIFAR-10, and ImageNet. (b) and (d) Corresponding images recovered by our reflection-mode single-fiber imaging technique.

Fig. 1(b). Although determining the amplitude-to-amplitude conversion, the reflection matrix R is complex valued. At the beginning, a complex-valued random matrix is assigned to it. For training purposes, a series of natural scene images (in our case, 50 000 images in 100 classes from the CIFAR-100 dataset[37]) are projected by a DMD as the ground truth X'. The single-fiber imaging system acquires the corresponding intensities of the resultant speckle patterns, from which we take only the amplitude Y. These two data are used as the training dataset. For each speckle pattern, the input amplitude is calculated with Eq. (1) as the output. After that, the loss function, $\zeta = \sum |X'-X|/N$, is calculated to assess the similarity metric between the outputs and the ground truth. We then apply a gradient descent-based iterative algorithm to optimize the parameters of R to reduce the loss function until the convergence of ζ to a minimum value. In practice, the optimization process is performed on the NVIDIA 1080 Ti using Keras/TensorFlow with the Adam optimizer. Learning to obtain a reflection matrix takes ~50 min. Finally, the optimized reflection matrix after training is used for image recovery from a single-shot intensity recording in terms of the object being imaged.

To test the capability of our proposed technique in imaging arbitrary natural scenes, we used the DMD to project different types of images that are not used as part of the training dataset. The images are recovered with the optimized reflection matrix from the corresponding output speckle patterns recorded at the proximal fiber end. Figure 2 shows the typical recovered images of the projected natural scenes together with the ground truth as a comparison. To quantify the quality of the image recovery, the PCCs are calculated, which is defined as

$$PCC = \frac{\sum (X - \overline{X})(X' - \overline{X'})}{\sqrt{\sum (X - \overline{X})^2 \sum (X' - \overline{X'})^2}}, \tag{2}$$

where $\overline{X}$ is the mean value of the grayscale image X. This coefficient measures the similarity between the ground truth and recovered

images with a maximum value of 1. Our technique achieves good performance in imaging different types of scenes. For example, the averaged PCCs for imaging both 7500 symbol scenes [see some pictures in Figs. 2(a) and 2(b)] from the MNIST[38] and Fashion-MNIST[39] datasets reach 0.86 and 0.88, respectively. For the more complex scenes with grayscale and detailed information from the CIFAR-10[37] and ImageNet[40] databases, we also obtain high PCCs of 0.87 and 0.84 in imaging 7500 scenes, as shown in Figs. 2(c) and 2(d). Overall, more than 30 000 scenes in total were imaged, and an averaged PCC over 0.84 with respect to the ground truth was achieved. These testing results suggest that the trained reflection matrix applies to the universal inverse conversion from the output amplitudes to the input amplitudes. Thus, combined with the proposed image reconstruction scheme, our technique is capable of imaging arbitrary scenes with high quality.

Benefiting from the ability of single-shot imaging, the single-fiber system also enables imaging dynamic scenes at high frame rates. In the experiment, the DMD projected videos at various frame rates, and the endoscope recorded the dynamic scenes. Figure 3 shows snapshots from movies of two dynamic scenes recovered by the endoscope. Figure 3(a) shows consecutive frames of a walking woman, and Fig. 3(b) shows the frames of a rotating clover. The full video is available in Video 1 (Multimedia view). We find that our imaging method is able to achieve high-quality recovery even though the dynamic scenes look similar. These two videos were recorded at a frame rate of 180 fps. The imaging speed is now limited by the acquisition rate of the camera. Thus, it is feasible to scale up the frame rate by equipping a higher-speed camera.

We further demonstrate endoscopic imaging with an enlarged FOV that is important for diagnostic endoscopy. For this purpose, we translated the distal end of the fiber to take images at different parts of the target object. First, a United States Air Force (USAF) resolution target projected on the DMD was imaged. The same trained reflection
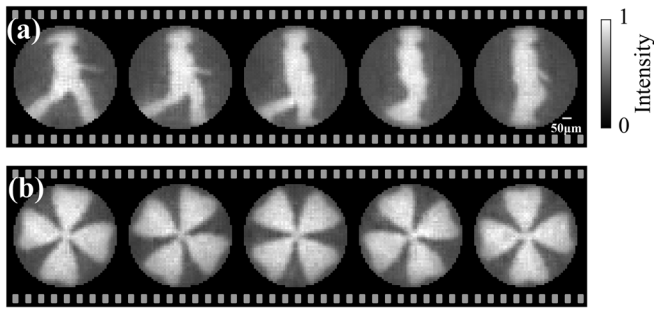
FIG. 3. Images of dynamic scenes captured by our single-fiber imaging technique. Snapshots from the movies showing a walking woman (a) and rotating clover (b). The videos were captured at a frame rate of 180 fps. The full video is available in Video 1. Multimedia view: https://doi.org/10.1063/5.0132123.1

matrix of the fiber was adopted for the image recovery from a single-shot speckle pattern for each position. The FOV of each single-shot image is ∼540 μm in diameter, and a larger image is obtained via a montage. Figure 4(a) shows the montage image of the USAF target with a size of ∼2.6 × 2.6 mm^2, where the features of group 3 and element 4 (dashed rectangle) can be clearly resolved. For further examination, the edge spread function of the square in the target was plotted in Fig. 4(b), and its corresponding line spread function (LSF) was fitted to compute the lateral resolution. The lateral resolution, defined by the full-width at half-maximum of the LSF, was 31.3 μm. In addition, to mimic the imaging of biological tissue, the fiber endoscope took a picture of a natural scene with blood vessels from a chicken embryo displayed by the DMD. The original picture (ground truth) is shown in
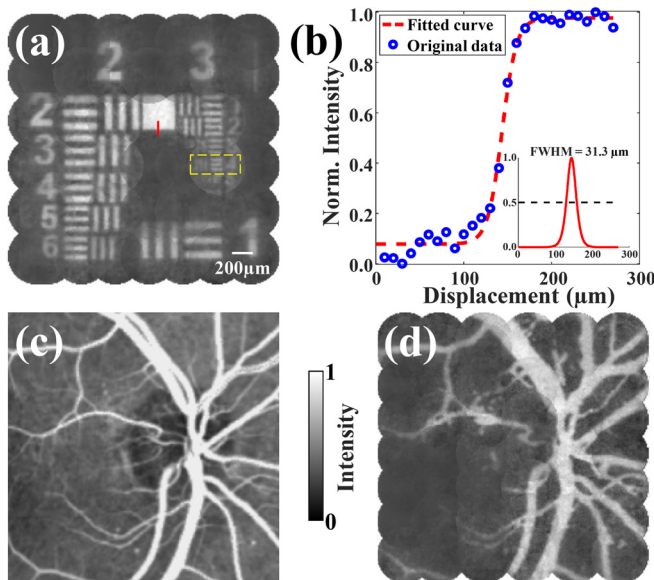

FIG. 4. Endoscopic imaging with an enlarged field of view. (a) The montage image of a USAF target. (b) Normalized edge spread function of the square in the target. The corresponding line spread function was fitted to compute the lateral resolution defined by the full-width at half-maximum (inset). (c) The original image of blood vessels from a chicken embryo. (d) The image recovered by our single-fiber imaging technique.

Fig. 4(c), and Fig. 4(d) presents the captured image with our method. Most of the structures can be clearly resolved, but some fine features are lost because of the limited resolution dictated by the DMD pixel pitch (10.8 μm in our case). If the conformation changes significantly due to the translation of the fiber, the reflection matrix would change. Thus, we need to re-calibrate the reflection matrix for the large-FOV endoscopic imaging.

To demonstrate real-time imaging of real objects, we applied the proposed technique to observe a moving microneedle that was used for microinjection, as illustrated in Fig. 5(a). In the experiment, the microneedle was freely moved in two dimensions, and the single-fiber imaging system recorded the dynamic process. Because the microneedle was almost transparent [see its bright-field image in Fig. 5(b)], the reflected light was slightly weak. Even so, the endoscope recovered its images from the recorded speckle patterns. Figures 5(c)–5(j) and a movie present the result of the dynamic scene with a video rate of 180 fps. The full video is available in Video 2 (Multimedia view). These results suggest that the reflection-mode single-fiber imaging scheme is expected to benefit the real-time observation of practical biomedical processes in a minimally invasive manner.

There are several ways to improve the imaging performance of our single MMF endoscope prototype. Unlike conventional deep learning-based imaging methods, our imaging scheme relies on learning the reflection matrix of the fiber. In practice, the training datasets were generated by the DMD. As a result, the spatial resolution was limited by the DMD. A DMD with a smaller pixel size and a higher pixel count can enhance the resolution. Note that the upper limit of the optical resolution is dictated by the NA of the MMF. In addition, the NA governs the light divergence from the fiber distal end, and a larger FOV is possible by increasing the distance from the fiber facet to the imaged object. However, this comes with a decrease in the spatial resolution. A montage strategy is now adopted to obtain a large FOV imaging with high resolution. Our imaging system has a depth of field of ∼200 μm, which can be further enhanced by ∼4.5 fold through a joint learning approach[10] (see Fig. S2 of the supplementary material). Currently, the frame rate of our single-shot imaging technique was mainly limited by the camera. Thus, equipping a higher-speed camera could further scale up the frame rate. With the current system, only the grayscale information of the target scenes was detected because a single-wavelength light source was used. The color information could also be detected. This can be achieved by building the reflection matrices for three primary colors (i.e., red, green, and blue) and recovering the grayscale image in each color.[20] We could extend the imaging capabilities of the single-fiber endoscope prototype to include depth information by the time-of-flight strategy.[11] Moreover, on-line calibration methods should be developed to avoid significant conformation change of the fiber for its practical use as an endoscopic probe.

In summary, we have demonstrated single-shot wide-field endoscopic imaging using a single MMF in reflection geometry. In such an endoscope, the single MMF simultaneously sends the illumination light and collects the backscattered light that is distorted in the fiber. To recover the object image from a single-shot output speckle pattern, a reflection matrix of the fiber was built by a learning-based approach for the universal inverse conversion from the output amplitudes to the input amplitudes under distorted illumination. To test the imaging performance of our technique, a DMD projected many arbitrary grayscale pictures with different complexities as the target scenes.
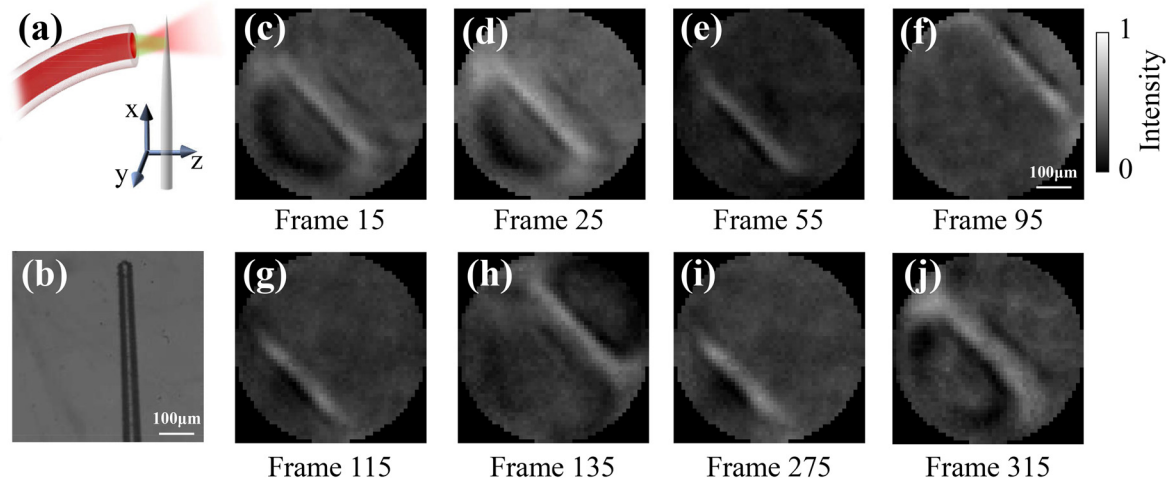
FIG. 5. Real-time observation of a freely moving microneedle. (a) Illustration of the imaging geometry. (b) Bright-field image of the microneedle. (c)–(j) Snapshots from the movie recording the dynamic process. The full video is available in Video 2. Multimedia view: https://doi.org/10.1063/5.0132123.2

Our technique imaged more than 30 000 natural scenes with an average PCC of over 0.84 in the experiment. In particular, the endoscope is also capable of imaging dynamic scenes at a frame rate of up to 180 fps and imaging with an enlarged FOV with the same setup. To demonstrate the ability of practical imaging, we applied the technique to image a freely moving microneedle, and real-time observation of the dynamic process was achieved. All these results demonstrate that video-rate wide-field endoscopic imaging in reflection can be realized by a simple and compact single-fiber lens-less system. Our work paves the way for practical single-fiber endoscopes with potential in biomedical and clinical studies.

See the supplementary material on experimental details and characterization of depth of field.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Yifan Liu: Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Visualization (lead); Writing – original draft (lead). **Panpan Yu:** Data curation (equal); Formal analysis (equal); Methodology (supporting); Validation (supporting). **Yijing Wu:** Methodology (equal); Writing – review & editing (supporting). **Ziqiang Wang:** Data curation (supporting); Methodology (equal); Writing – review & editing (supporting). **Yin Mei Li:** Conceptualization (supporting); Funding acquisition (equal); Writing – review & editing (equal). **Jinyang Liang:** Conceptualization (supporting); Writing – review & editing (equal). **Puxiang Lai:** Funding acquisition (supporting); Writing – review & editing (equal). **Lei Gong:** Conceptualization (lead); Funding acquisition (lead); Project administration (lead); Supervision (lead); Validation (equal); Writing – original draft (equal); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, Biomed. Opt. Express **4**(2), 260 (2013).

[2] S. Turtaev, I. T. Leite, T. Altwegg-Boussac, J. M. P. Pakan, N. L. Rochefort, and T. Cizmar, Light **7**, 92 (2018).

[3] T. Zhong, Z. Qiu, Y. Wu, J. Guo, H. Li, Z. Yu, S. Cheng, Y. Zhou, J. Zhu, J. Tian, L. Sun, and P. Lai, Adv. Photonics Res. **3**, 2100231 (2021).

[4] S. A. Vasquez-Lopez, R. Turcotte, V. Koren, M. Plöschner, Z. Padamsey, M. J. Booth, T. Čižmár, and N. J. Emptage, Light **7**(1), 110 (2018).

[5] J. Shin, D. N. Tran, J. R. Stroud, S. Chin, T. D. Tran, and M. A. Foster, Sci. Adv. **5**(12), eaaw5595 (2019).

[6] B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. M. Cheung, and M. J. Schnitzer, Nat. Methods **2**(12), 941 (2005).

[7] A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, Opt. Express **24**(15), 16835 (2016).

[8] W. Göbel, J. N. D. Kerr, A. Nimmerjahn, and F. Helmchen, Opt. Lett. **29**(21), 2521 (2004).

[9] Z. Ju, Z. Yu, Z. Meng, N. Zhan, L. Gui, and K. Xu, Opt. Express **30**(9), 15596 (2022).

[10] L. Wang, Y. Yang, Z. Liu, J. Tian, Y. Meng, T. Qi, T. He, D. Li, P. Yan, M. Gong, Q. Liu, and Q. Xiao, Laser Photonics Rev. **16**, 2100724 (2022).

[11] D. Stellinga, D. B. Phillips, S. P. Mekhail, A. Selyem, S. Turtaev, T. Čižmár, and M. J. Padgett, Science **374**(6573), 1395 (2021).

[12] I. T. Leite, S. Turtaev, D. E. B. Flaes, and T. Čižmár, APL Photonics **6**(3), 036112 (2021).

[13] Q. Zhao, P. Yu, Y. Liu, Z. Wang, Y. Li, and L. Gong, Appl. Phys. Lett. **116**(18), 181101 (2020).

[14] S. Li, S. A. R. Horsley, T. Tyc, T. Cizmar, and D. B. Phillips, Nat. Commun. **12**(1), 3751 (2021).

[15] P. Fan, M. Ruddlesden, Y. Wang, L. Zhao, C. Lu, and L. Su, Laser Photonics Rev. **15**(4), 2000348 (2021).

[16] S. Resisi, S. M. Popoff, and Y. Bromberg, Laser Photonics Rev. **15**(10), 2000553 (2021).

[17] C. Zhu, E. A. Chan, Y. Wang, W. Peng, R. Guo, B. Zhang, C. Soci, and Y. Chong, Sci. Rep. **11**(1), 896 (2021).

[18] T. Čižmár and K. Dholakia, Nat. Commun. **3**(1), 1027 (2012).

[19] N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, Optica **5**(8), 960 (2018).

[20] P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, Nat. Commun. **10**(1), 2029 (2019).

[21] Z. Yu, H. Li, T. Zhong, J. Park, S. Cheng, C. M. Woo, Q. Zhao, J. Yao, Y. Zhou, X. Huang, W. Pang, H. Yoon, Y. Shen, H. Liu, Y. Zheng, Y. Park, L. V. Wang, and P. Lai, Innovation **3**(5), 100292 (2022).

[22] S. Cheng, T. Zhong, C. M. Woo, Q. Zhao, H. Hui, and P. Lai, Opt. Express **30**(18), 32565 (2022).

[23] Z. Wen, L. Wang, X. Zhang, Y. Ma, X. Liu, C. F. Kaminski, and Q. Yang, Opt. Lett. **45**(17), 4931 (2020).

[24] Z. Dong, Z. Wen, C. Pang, L. Wang, L. Wu, X. Liu, and Q. Yang, Sci. Bull. **67**, 1224 (2022).

[25] Z. Liu, L. Wang, Y. Meng, T. He, S. He, Y. Yang, L. Wang, J. Tian, D. Li, P. Yan, M. Gong, Q. Liu, and Q. Xiao, Nat. Commun. **13**(1), 1433 (2022).

[26] L. V. Amitonova and J. F. de Boer, Light **9**, 81 (2020).

[27] B. Lochocki, M. V. Verweg, J. J. M. Hoozemans, J. F. de Boer, and L. V. Amitonova, APL Photonics **7**(7), 071301 (2022).

[28] M. Plöschner, T. Tyc, and T. Čižmár, Nat. Photonics **9**, 529 (2015).

[29] L. V. Amitonova and J. F. de Boer, Opt. Lett. **43**(21), 5427 (2018).

[30] A. M. Caravaca-Aguirre, S. Singh, S. Labouesse, M. V. Baratta, R. Piestun, and E. Bossy, APL Photonics **4**(9), 096103 (2019).

[31] Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, Phys. Rev. Lett. **109**(20), 203901 (2012).

[32] S. Lee, V. J. Parot, B. E. Bouma, and M. Villiger, Optica **9**(1), 112 (2022).

[33] Y. Liu, Z. Zhang, P. Yu, Y. Wu, Z. Wang, Y. Li, W. Liu, and L. Gong, Appl. Phys. Lett. **120**(13), 131101 (2022).

[34] G. S. D. Gordon, M. Gataric, A. G. C. P. Ramos, R. Mouthaan, C. Williams, J. Yoon, T. D. Wilkinson, and S. E. Bohndiek, Phys. Rev. X **9**(4), 041050 (2019).

[35] P. Fan, T. Zhao, and L. Su, Opt. Express **27**(15), 20241 (2019).

[36] M. W. Matthès, Y. Bromberg, J. d Rosny, and S. M. Popoff, Phys. Rev. X **11**(2), 021060 (2021).

[37] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Report No. TR-2009 (University of Toronto, 2009).

[38] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Proc. IEEE **86**(11), 2278 (1998).

[39] H. Xiao, K. Rasul, and R. Vollgraf, arXiv:1708.07747 (2017).

[40] J. Deng, W. Dong, R. Socher, L. J. Li, L. Kai, and F.-F. Li, paper presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009.