

Temporal compressive super-resolution microscopy at frame rate of 1200 frames per second and spatial resolution of 100 nm

Yilin He,^{a,†} Yunhua Yao,^{a,†} Dalong Qi,^a Yu He,^a Zhengqi Huang,^a Pengpeng Ding,^a Chengzhi Jin,^a Chonglei Zhang,^b Lianzhong Deng,^a Kebin Shi,^c Zhenrong Sun,^a Xiaocong Yuan,^{b,*} and Shian Zhang^{a,d,*}

^aEast China Normal University, School of Physics and Electronic Science, State Key Laboratory of Precision Spectroscopy, Shanghai, China

^bShenzhen University, Institute of Microscale Optoelectronics, Nanophotonics Research Center, Shenzhen Key Laboratory of Micro-Scale Optical Information Technology, Shenzhen, China

^cPeking University, School of Physics, Frontiers Science Center for Nanooptoelectronics, State Key Laboratory for Mesoscopic Physics, Beijing, China

^dShanxi University, Collaborative Innovation Center of Extreme Optics, Taiyuan, China

Abstract. Various super-resolution microscopy techniques have been presented to explore fine structures of biological specimens. However, the super-resolution capability is often achieved at the expense of reducing imaging speed by either point scanning or multiframe computation. The contradiction between spatial resolution and imaging speed seriously hampers the observation of high-speed dynamics of fine structures. To overcome this contradiction, here we propose and demonstrate a temporal compressive super-resolution microscopy (TCSR) technique. This technique is to merge an enhanced temporal compressive microscopy and a deep-learning-based super-resolution image reconstruction, where the enhanced temporal compressive microscopy is utilized to improve the imaging speed, and the deep-learning-based super-resolution image reconstruction is used to realize the resolution enhancement. The high-speed super-resolution imaging ability of TCSR with a frame rate of 1200 frames per second (fps) and spatial resolution of 100 nm is experimentally demonstrated by capturing the flowing fluorescent beads in microfluidic chip. Given the outstanding imaging performance with high-speed super-resolution, TCSR provides a desired tool for the studies of high-speed dynamical behaviors in fine structures, especially in the biomedical field.

Keywords: super-resolution microscopy; high-speed imaging; compressive sensing; deep learning; image reconstruction.

Received Nov. 16, 2022; revised manuscript received Feb. 16, 2023; accepted for publication Feb. 21, 2023; published online Mar. 9, 2023.

© The Authors. Published by SPIE and CLP under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.

[DOI: [10.1117/1.AP.5.2.026003](https://doi.org/10.1117/1.AP.5.2.026003)]

1 Introduction

Exploring fine structures and their dynamics beyond the optical diffraction limit is an urgent requirement in many research fields, especially in biology and medicine. To date, various super-resolution microscopy techniques have been developed to surpass the optical diffraction limit. For example, stimulated emission depletion microscopy (STED) improved the resolution

by shrinking the point spread function (PSF) with nonlinear stimulated emission depletion based on confocal microscopy.¹ Single molecule localization microscopy (SMLM), involving photoactivated localization microscopy (PALM)² and stochastic optical reconstruction microscopy (STORM),³ achieved a higher resolution by localizing a single fluorescent molecule with sparse fluorescence activation instead of recording fluorescence distribution. Structured illumination microscopy (SIM) obtained a super-resolution image by loading normally inaccessible high spatial frequency information into the recorded images by the moiré effect.⁴ Super-resolution optical fluctuation imaging (SOFI) utilized random temporal signal fluctuations of single emitters

*Address all correspondence to Xiaocong Yuan, xyuan@szu.edu.cn; Shian Zhang, sazhang@phy.ecnu.edu.cn

[†]The authors contributed equally to this work.

to achieve background-free super-resolution microscopy based on high-order statistics.⁵ In addition, some novel microscopy techniques are emerging by combining multiple super-resolution imaging methods. For example, a combining method with STED and SMLM realized better resolution and less fluorophore bleaching, such as minimal STED⁶ or minimal photon fluxes (MINFLUX).⁷ A STED-SIM method achieved 30 nm resolution and single-molecule sensitivity by utilizing STED to provide nonlinear modulation for SIM.⁸ A SIM-based point localization estimator (SIMPLE) method obtained simultaneous particle localization with twofold precision by using phase-shifted sinusoidal wave patterns as nanometric rulers.⁹ Super-resolution microscopy, as a powerful imaging tool, has boosted the development of biomedicine, and numerous discoveries have been reported,¹⁰ such as centrosome structure and function,^{11,12} nuclear and chromatin organization,^{13,14} and mitochondrial membrane protein organization.¹⁵

It should be noted that all the techniques mentioned above acquire the super-resolution ability at the expense of reducing the imaging speed by either point scanning or multiframe computation. Thus, the imaging speed is inevitably limited, which greatly affects the observation of high-speed dynamics of fine structures. Recently, a single-image super-resolution (SISR) technique was proposed to overcome the limited imaging speed by extracting a super-resolution image from one recorded image, which allowed the super-resolution imaging speed to reach the frame rate of a camera. Many deep-learning-based algorithms with neural networks have accelerated the development of SISR due to their outstanding image processing ability. For example, Wang et al.¹⁶ employed a generative adversarial network (GAN) to realize cross-modality super-resolution from confocal microscopy images to STED images or from total internal reflection fluorescence (TIRF) images to SIM images. Chen et al.¹⁷ proposed a novel network combining a super-resolution network and a signal-enhancement network to transfer wide-field images to SMLM images. Qiao et al.¹⁸ developed a deep Fourier channel attention network (DFCAN) for super-resolution imaging by leveraging the frequency content difference across distinct features to learn precise hierarchical representations of high-frequency information in diverse biological structures. Obviously, SISR improves the super-resolution imaging speed by avoiding the point scanning and multiframe computation, but the imaging speed is still restricted by the frame rate of a camera.

To further improve the super-resolution imaging speed that breaks through the frame rate limit of a camera, we propose and demonstrate a novel temporal compressive super-resolution microscopy technique, termed TCSR, which combines an enhanced temporal compressive microscopy and a deep-learning-based image reconstruction. Here, the purpose of the enhanced temporal compressive microscopy is to improve the imaging speed by reconstructing multiple images from one compressed image, and the deep-learning-based image reconstruction seeks to achieve the super-resolution without reduction in the imaging speed. The high-speed super-resolution imaging ability of TCSR is verified in theory and experiment, and the experimental result shows that TCSR has the imaging capability with a frame rate of 1200 frames per second (fps) and spatial resolution of 100 nm based on a 200 fps CMOS and a 100× objective lens. TCSR can provide a well-established tool for capturing the high-speed dynamics of fine structures and will have promising applications in the biomedical field.

2 Theoretical Model

As an inherent feature, natural dynamic scenes have sparsity in some transform domains. Thus, the spatiotemporal information of a dynamic scene can be recovered from a compressed sampling based on compressive sensing theory.^{19,20} Moreover, the spatial distributions at adjacent moments have continuity. Therefore, the spatial distribution at a moment can provide the reference information for the dynamic scene.^{21,22} Based on these premises, we propose an enhanced temporal compressive microscopy to capture the high-speed dynamic scene, which combines the spatiotemporal compressive information and the transient spatial information. The imaging model is shown in Fig. 1(a). The original dynamic scene $D(x, y, t)$ is first transferred into a diffraction-limited dynamic scene $B(x, y, t)$ after passing through an optical microscope. This process can be treated as the convolution with the PSF of the microscope, and is expressed as

$$B(x, y, t) = \mathbf{H}D(x, y, t) = D(x, y, t) * \text{PSF}(x, y), \quad (1)$$

where \mathbf{H} is the diffraction limitation operator, and $\text{PSF}(x, y)$ is the PSF of the microscope. The diffraction-limited dynamic scene is then synchronously sampled by two channels: a compressive sampling (CS) channel and a transient sampling (TS) channel. The CS channel is utilized to collect all the spatiotemporal information of the diffraction-limited dynamic scene, while the TS channel is used to acquire the spatial information at a moment of this dynamic scene. In the CS channel, the diffraction-limited dynamic scene $B(x, y, t)$ is sequentially encoded by a programmable spatial light modulator with random patterns $C(x, y, t)$, and then the coded scene is recorded as a compressed image $M_{\text{cs}}(x, y)$ by a camera with a long exposure time Δt_{cs} . The compressed image in the CS channel can be formulated as

$$M_{\text{cs}}(x, y) = \mathbf{O}_{\text{cs}}B(x, y, t) = \int_0^{\Delta t_{\text{cs}}} B(x, y, t) \times C(x, y, t) dt, \quad (2)$$

where \mathbf{O}_{cs} represents the CS operator. In the TS channel, a transient fragment in the diffraction-limited dynamic scene $B(x, y, t)$ is captured by a camera with a short exposure time Δt_{ref} as reference. The reference image $M_{\text{ref}}(x, y)$ in the TS channel can be expressed as

$$M_{\text{ref}}(x, y) = \mathbf{O}_{\text{ref}}B(x, y, t) = \int_{t_0}^{t_0 + \Delta t_{\text{ref}}} B(x, y, t) dt, \quad (3)$$

where \mathbf{O}_{ref} represents the TS operator. Because $\Delta t_{\text{cs}} \gg \Delta t_{\text{ref}}$, the reference image $M_{\text{ref}}(x, y)$ in the TS channel can be treated as a transient frame in the dynamic scene $B(x, y, t)$. Based on the spatiotemporal continuity within the dynamic scene, $B(x, y, t)$ can be further written as

$$\begin{aligned} B(x, y, t) &= \mathbf{R}(M_{\text{ref}}(x, y), V) \\ &= M_{\text{ref}}(x, y) + \int_0^t \partial_{\tau} B(x, y, \tau) |_{\tau=0} d\tau \\ &= M_{\text{ref}}(x, y) + \int_0^t \left(\frac{\partial B}{\partial x} \frac{\partial x}{\partial \tau} + \frac{\partial B}{\partial y} \frac{\partial y}{\partial \tau} + \frac{\partial B}{\partial \tau} \right) \Big|_{\tau=0} d\tau \\ &\approx M_{\text{ref}}(x, y) + \int_0^t (\partial_x M_{\text{ref}} \cdot V_x + \partial_y M_{\text{ref}} \cdot V_y) d\tau, \quad (4) \end{aligned}$$

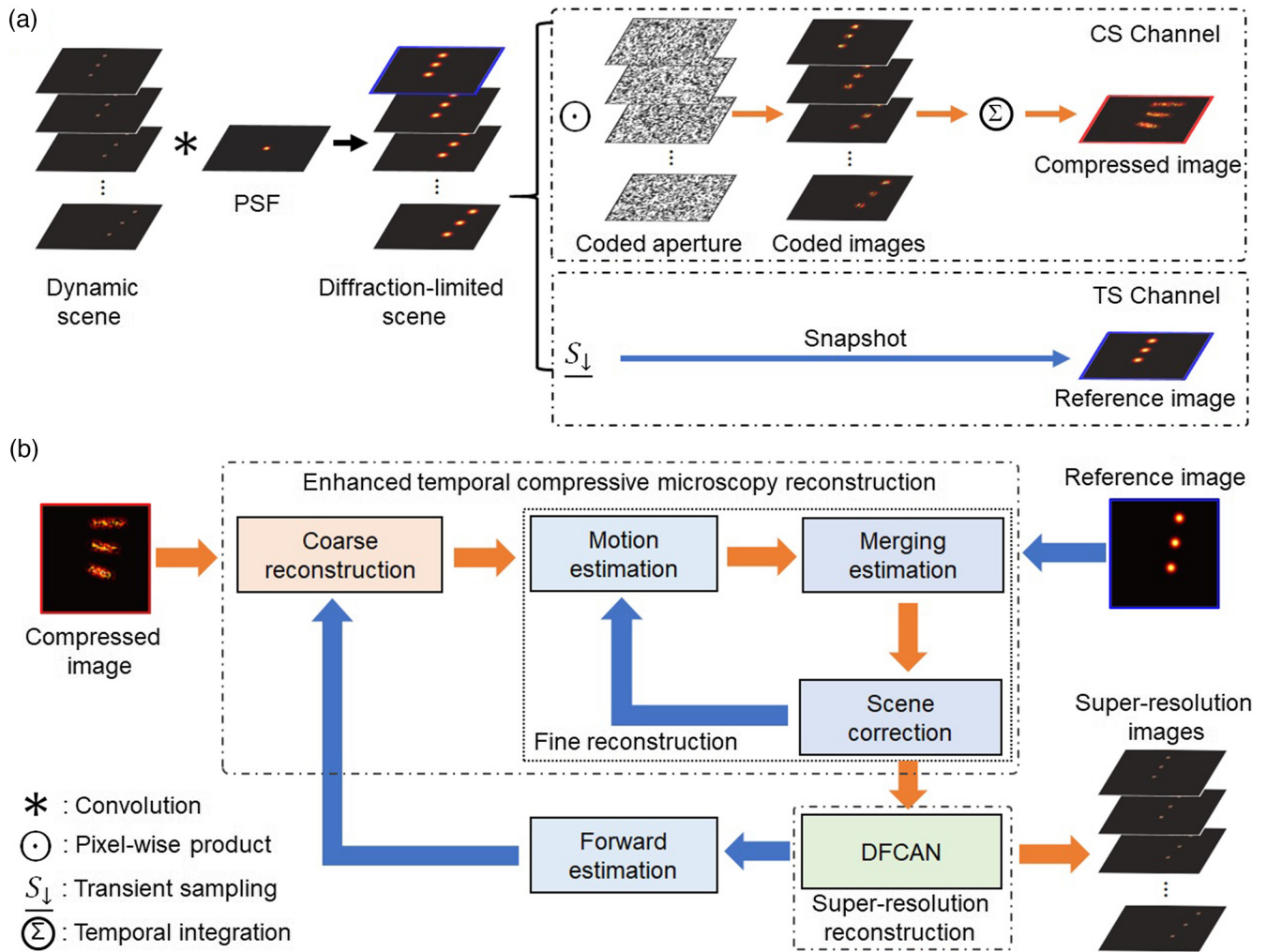


Fig. 1 Theoretical model of TCSR. (a) Image acquisition flowchart of TCSR. (b) Image reconstruction framework of TCSR.

where \mathbf{R} is the merging estimation operator, and V_x and V_y are the components of the motion vector V in the horizontal and vertical directions. Based on the compressed sensing theory, one can get the estimation of the dynamic scene by solving a constrained optimization problem, which is given as

$$D = \arg \min \|\mathbf{O}_{cs}B - M_{cs}\|_2^2 + \lambda \|\mathbf{O}_{ref}B - M_{ref}\|_2^2 + \rho \|\Phi D\|_0$$

$$\text{subject to } \begin{cases} B = \mathbf{R}(M_{ref}, V) \\ B = HD \end{cases}, \quad (5)$$

where λ is the channel weight factor that is determined by the light flux ratio between the CS and TS channels, ρ is the regularization factor of sparsity constraint, and Φ is the sparse transform operator. For simplicity, the indices in Eq. (5) are omitted. To solve this constrained optimization problem, Eq. (5) is split into three subproblems, involving a compressed sensing recovery problem, a motion estimation problem, and an image super-resolution problem. The three subproblems are solved step by step by alternative iteration with constraints,²³ which can be expressed as

$$\text{Step 1: } B_{cr} = \arg \min_B \|\mathbf{O}_{cs}B - M_{cs}\|_2^2 + \rho_1 \|\Phi_1 B\|_0, \quad (6)$$

$$\text{Step 2: } B_{fr}^{(n)} = \arg \min_B \|\mathbf{O}_{cs}B - M_{cs}\|_2^2 + \lambda \|\mathbf{O}_{ref}B - M_{ref}\|_2^2 + \rho_2 \|\Phi_2 B\|_0$$

$$\text{subject to } B = \mathbf{R}(M_{ref}, \mathbf{F}(B_{fr}^{(n-1)})), \quad (7)$$

$$\text{Step 3: } D_r^{(m)} = \mathbf{S}(B_{fr}), \quad (8)$$

where B_{cr} is the coarse reconstruction result of the diffraction-limited dynamic scene based on the compressed image M_{cs} from the CS channel, B_{fr} is the fine reconstruction result combining the spatiotemporal information from the CS and TS channels, D_r is the recovered super-resolution dynamic scene, ρ_1 and ρ_2 are the regularization parameters in each subproblem, Φ_1 and Φ_2 are the sparse transform operators in each subproblem, \mathbf{F} is the motion vector estimation operator, \mathbf{S} is the SISR operator, and n and m are the iteration numbers in Eqs. (7) and (8), respectively. Using a gradient descent method,²⁴ Eqs. (6)–(8) can be iteratively calculated by considering the balance among super-resolution mapping, measurement constraint, and sparsity constraint.

The image reconstruction framework of TCSR is shown in Fig. 1(b). The compressed image in the CS channel is first

recovered to a diffraction-limited dynamic scene B_{cr} with a plug-and-play (PnP) algorithm embedded with multiple denoisers, including total variation (TV),¹⁹ FFDnet,²⁵ and FastDVDnet.²⁶ Here various priors of images and videos are utilized. The reconstruction from the compressed image can only obtain the coarse spatiotemporal information of the dynamic scene due to the insufficient sampling. The result is further processed by a fine reconstruction, containing motion estimation, merging estimation, and scene correction modules. The motion estimation is conducted to extract the dynamic features. In this process, the motion vector V is determined by a sparse motion estimation algorithm based on block-matching²⁷ and Lucas–Kanade optical flow.²⁸ Then, the reference image M_{ref} from the TS channel and the motion vector V are fused in the merging estimation module to provide an estimation for the dynamic scene by combining the spatiotemporal information from the two channels. The errors in the motion estimation and merging estimation are compensated in the scene correction module by optimizing the details of the dynamic scene based on the PnP algorithm with the measurement and prior constraints. The three modules are conducted iteratively to acquire the final dynamic scene with fine details, which satisfies the measurement constraints, motion estimation, and prior constraints simultaneously. The reconstructed dynamic scene is further processed by a super-resolution reconstruction module. Here, a pretrained DFCAN is utilized to handle the task, which is a residual network with Fourier channel attention blocks. Exploiting the power spectrum characteristics of distinct feature maps in the Fourier domain, DFCAN can bridge the low-resolution and high-resolution image spaces precisely.¹⁸ A forward estimation is used to calculate the error E between the reconstruction results and actual measurements, involving the compressed image and reference image, which is expressed as

$$E^{(m)} = \|\mathbf{O}_{cs} \mathbf{H} \mathbf{D}_r^{(m)} - M_{cs}\|_2^2 + \lambda \|\mathbf{O}_{ref} \mathbf{H} \mathbf{D}_r^{(m)} - M_{ref}\|_2^2. \quad (9)$$

Here, the error will be calculated in each iteration. Once the error reaches the preset threshold, the desired super-resolution dynamic scene is obtained. It is worth mentioning that TCSR is rather different from the simple concatenation of temporal compressive imaging and DFCAN. The recovered images from temporal compressive imaging have different features compared with natural images, which is due to the information loss during compressive acquisition and the imperfect optimization during image reconstruction. Moreover, the DFCAN is trained by using natural image pairs with low and high resolution. The mismatch in image features makes it difficult to obtain acceptable results because of the generalization problem in end-to-end networks. However, TCSR utilizes the additional reference frame to recover the images with higher accuracy, which decreases the mismatch in image features between recovered images and natural images. In addition, the global iterations of compressive image reconstruction and super-resolution processing are conducted in TCSR to optimize the final super-resolution images with corresponding forward estimation. In this way, the super-resolution ability of DFCAN can be fully utilized, alleviating the generalization problem.

3 Simulation Result

In order to verify the feasibility of TCSR, we design a dynamic scene with high-speed moving nanorings for simulation. In the simulation, the diameter of the rings is 750 nm, and the

width of the rings (full width at half-maximum, FWHM) is 96 nm. The three nanorings move in different ways. The top one moves right with a constant velocity of 97.5 nm/frame, the middle one moves right with an initial velocity of 32.5 nm/frame and a rightward acceleration of 3.25 nm/frame², and the bottom one moves along a curve with an initial velocity of 91.91 nm/frame with direction 45 deg to the horizontal and a rightward acceleration of 3.25 nm/frame². The fluorescence wavelength of the nanorings is 560 nm, and the numerical aperture (NA) of the objective lens is 1.5. The dynamic scene contains 36 images with the size of 512 × 512, which is utilized as the ground truth (GT). After passing through a microscope, the diffraction-limited dynamic scene is individually sampled by the CS and TS channels. Six compressed images with a size of 128 × 128 are acquired in the CS channel. Thus, the data compression ratio is 6, which means that one compressed image contains the information of six original images. However, six reference images with a size of 256 × 256 are recorded in the TS channel, which means that one of six original images is selected to provide a reference for the image reconstruction in the CS channel. The images from the two channels are then processed by the reconstruction algorithm in Fig. 1(b) to recover the original dynamic scene. One compressed image is shown in Fig. 2(a), together with the corresponding reference image. As can be seen, the compressed image shows an obvious blur due to spatial coding and temporal integration, while the reference image shows a clear profile of the nanorings. The reconstructed images by TCSR are shown in Fig. 2(b), associated with the GT images. Similarly, the reconstructed images have a clear spatial profile. The width of the rings is much smaller than that in the reference image, and it is close to that in the GT images. The average peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of the recovered images by TCSR are 27.87 dB and 0.98, respectively. The motion traces of the nanorings extracted from TCSR are given in Fig. 2(c), associated with the motion traces from GT. Obviously, the motion traces of TCSR agree well with those of GT, which demonstrates that TCSR can recover the dynamic scene containing the objects with various displacement patterns with high temporal accuracy. For quantitatively characterizing the effect of super-resolution, the radial intensity distributions of the nanorings in the reference, GT, and TCSR images are extracted, as shown in Fig. 2(d). The width of the nanorings is decreased to 102 nm in the TCSR images from 251 nm in the reference image, which is close to 96 nm in the GT images. The whole video of the high-speed moving nanorings is provided in Video 1.

4 Experimental Design

The experimental arrangement of TCSR is shown in Fig. 3. A continuous-wave laser with the wavelength of 532 nm (Laser Quantum, Torus 532) is used as the excitation source. The laser beam is expanded by a beam expander and reflected by a dichroic mirror and is then focused in the microchannel of a customized glass microfluidic chip on the sample stage with an objective lens (Olympus, UPlanApo, Oil, 100×, NA 1.5). The depth and width of the microchannel are 10 and 120 μm, respectively. Fluorescent beads (ThermoFisher, F8800) with the diameters of about 100 nm and an emission wavelength of 560 nm are dispersed in distilled water and then injected into the microchannel by an injection pump (MesoBioSys, MS-102P) with adjustable flow rate. The fluorescence signal passes through the dichroic mirror and then is divided into two

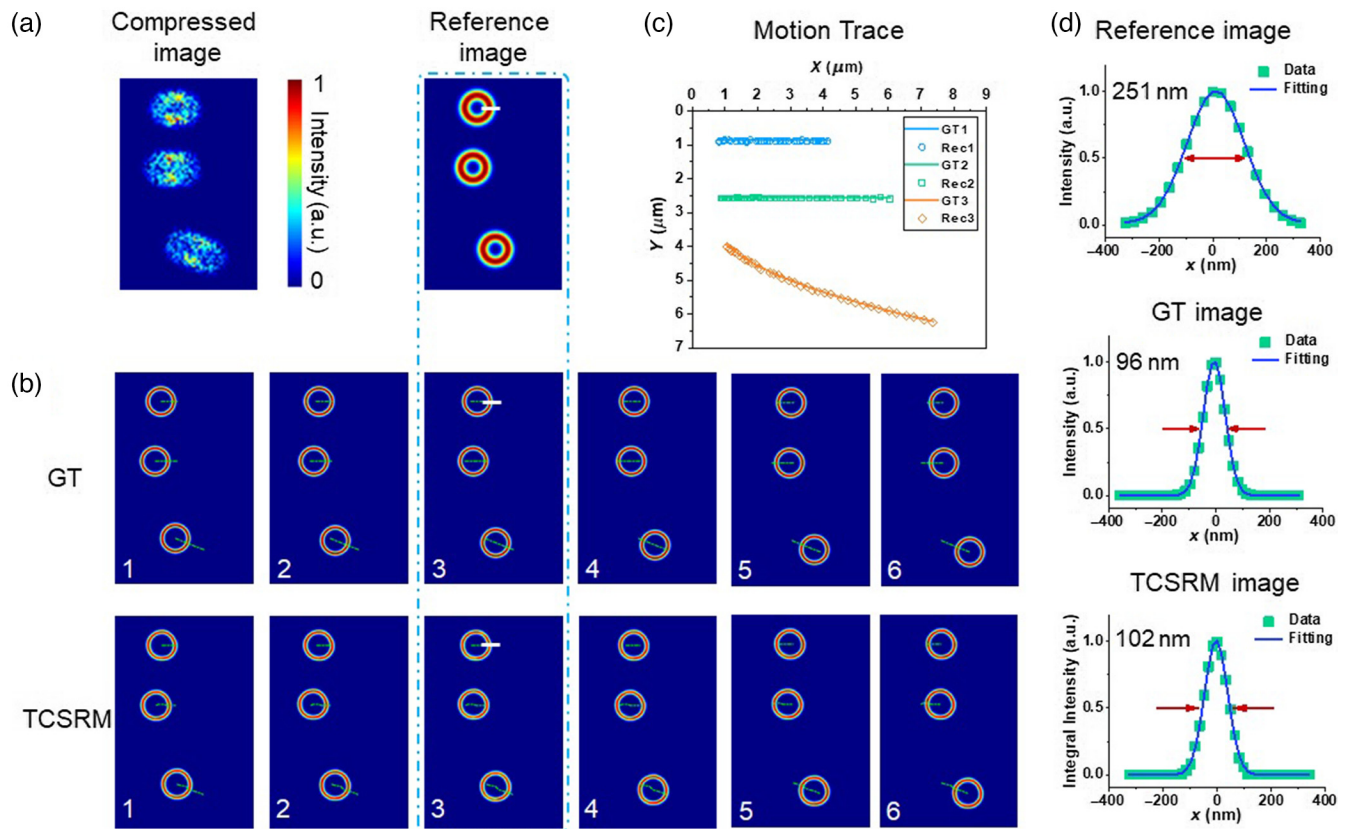


Fig. 2 Simulation result of moving nanorings by TCSR. (a) Compressed image and reference image measured by two channels in TCSR. (b) GT and TCSR images for six consecutive frames. The moving trajectories of the nanorings are labeled with green lines. (c) Motion traces of the three nanorings in the whole scene from GT (lines) and reconstructed result by TCSR (circles, squares, and rhombuses). (d) Radial intensity distributions of the nanorings along the white line in the reference, GT, and TCSR images (Video 1, mp4, 845 KB [URL: <https://doi.org/10.1117/1.AP.5.2.026003.s1>]).

components by a beam splitter: one is imaged on a digital micro-mirror device (DMD, Texas Instrument, DLP6500) for spatio-temporal encoding, and then recorded by a camera CMOS1 (Andor, Zyla 5.5); the other is directly recorded by a camera CMOS2 (Andor, Zyla 5.5). The DMD has a micromirror array of 1920×1080 with a size of $7.56 \mu\text{m}$. Here, CMOS1 is utilized to acquire compressed images of the dynamic scene, and CMOS2 is used to acquire transient images at some specific moments of the dynamic scene. A field programmable gate array (FPGA) device provides the trigger signals to synchronize the cameras and DMD accurately. The time sequences of these devices are shown in the inset of Fig. 3. The frame rates of both cameras are set as 200 fps, while the refresh rate of DMD is set as 1200 Hz. The exposure times of CMOS1 and CMOS2 are set to be 4.9 and 0.7 ms, respectively. In each exposure, the compressed image by CMOS1 contains the information of the dynamic scene with six spatial encodings, and the reference image by CMOS2 only records the transient information at a moment of the dynamic scene. In the reconstruction for the experimental data, the regularization parameters ρ_1 and ρ_2 are both set as 0.07, and the iteration numbers for coarse reconstruction, fine reconstruction, and forward estimation are 224, 64, and 4, respectively. The pixel number of the reconstructed images is 1024×2048 .

5 Experimental Result

The experimental result of a flowing fluorescent bead in a microfluidic chip is shown in Fig. 4, and the whole video is provided in Video 2. One selected compressed image and corresponding reference image are given in Fig. 4(a), and the reconstructed six images by TCSR are shown in Fig. 4(b). The size of the bead in TCSR is obviously decreased compared with that in the reference image, and the moving trajectory can be clearly distinguished. To show the resolution improvement, the intensity distributions of the bead along the horizontal and vertical directions in the reference image and the first frame of TCSR images are extracted and given in Figs. 4(c) and 4(d). The sizes of the bead (FWHM) in the horizontal and vertical directions for the reference image are 264 and 237 nm, and those in the TCSR image are 118 and 93 nm, respectively. Thus, the resolution is improved by a factor of about 2.2. That is to say, TCSR has the high-speed super-resolution ability with the frame rate of 1200 fps and the spatial resolution of about 100 nm, which surpasses conventional microscopy. The difference in the sizes in the two directions is due to the high-speed moving of the bead in the horizontal direction, which results in the stretch of the bead in this direction during the image reconstruction. According to the measurement of TCSR, the

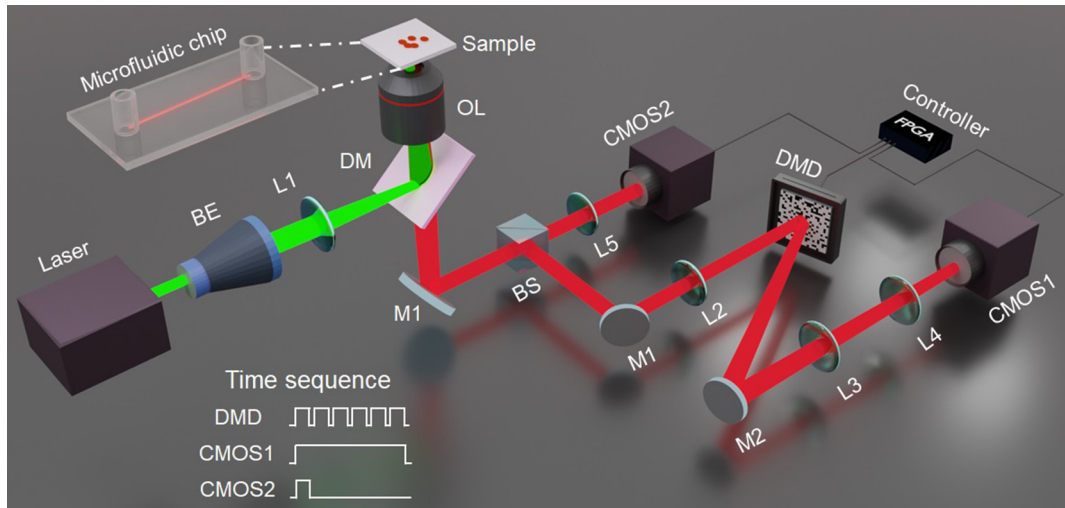


Fig. 3 Experimental design of TCSR. BE: beam expander; L: lens; DM: dichromatic mirror; OL: objective lens; DMD: digital micromirror device.

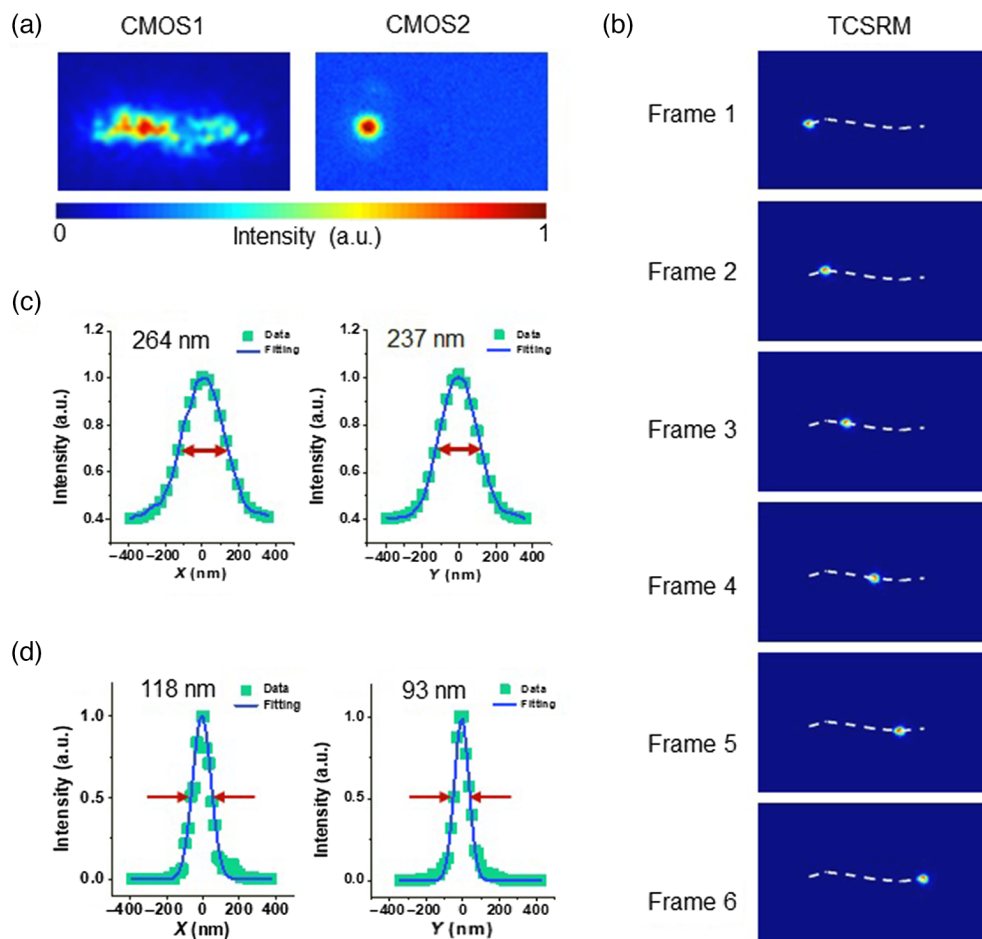


Fig. 4 Experimental result of flowing fluorescent bead in microchannel by TCSR. (a) Compressed and reference images recorded by two cameras. (b) Reconstructed images by TCSR. The trajectory of the moving bead is marked with white dashed lines. (c) and (d) Intensity distributions of the fluorescent bead along the horizontal and vertical directions in the reference image and the first frame in TCSR images (Video 2, mp4, 89.7 KB [URL: <https://doi.org/10.1117/1.AP.5.2.026003.s2>]).

average speed of the bead is about 0.39 mm/s, which is close to the flowing speed of the water with 0.42 mm/s. Here, the flowing speed is calculated based on the flux of the water and the size of the microchannel. Moreover, the bead does not move along a straight line, which may result from turbulent flow²⁹ or Brownian movement.³⁰ By measuring the speeds of the fluorescent beads at different locations with TCSR, the flowing speed distribution of the microchannel can also be extracted.

6 Discussion and Conclusion

TCSR is a lossy imaging by spatial encoding, which will reduce the image quality. One way is to improve the sampling rate in hardware, such as multiple CS channels, and the other way is to develop a more advanced image reconstruction algorithm in software, such as hybrid super-resolution algorithm. The reference image in the TS channel provides detailed spatial information for the image reconstruction in the CS channel, and therefore the exposure time of CMOS2 should be as short as possible under the condition of ensuring high enough signal-to-noise ratio. In general, the maximum exposure time should be shorter than the division of the exposure time of CMOS1 and the compressive ratio. Additionally, in order to obtain the effect of the super-resolution, the data compression ratio in the CS channel cannot be too large, and the value of around 10 is appropriate. An important application of TCSR is biomedical imaging. Compared with other wide-field super-resolution imaging, such as SIM and SISR, TCSR has lower light flux due to two-channel sampling, and the light field in the CS channel is spatially modulated in amplitude, while that in the TS channel is partially detected in a very short time scale. An end-to-end deep-learning super-resolution algorithm is utilized in TCSR, which has the limitation in generalization. Transfer learning³¹ can be used to reduce the required training datasets. Meanwhile, self-supervised networks, such as GAN³² and deep image prior,³³ may be adopted to improve the generalization of TCSR.

In conclusion, we have developed a high-speed super-resolution microscopy technique TCSR by combining an enhanced temporal compressive microscopy and a deep-learning-based image reconstruction. The enhanced temporal compressive microscopy realizes the high-speed imaging and the deep-learning-based image reconstruction obtains the resolution beyond the optical diffraction limit. Both the theoretical and experimental results verify the high-speed super-resolution imaging ability of TCSR, and the imaging performance with a frame rate of 1200 fps and spatial resolution of 100 nm is experimentally obtained. TCSR provides a powerful tool for the observation of high-speed dynamics of fine structures, especially in hydro-mechanics and biomedical fields, such as microflow velocity measurement,³⁴ organelle interactions,³⁵ intracellular transports,³⁶ and neural dynamics.³⁷ In addition, the framework of TCSR can also offer guidance for achieving higher imaging speed and spatial resolution in holography,³⁸ coherent diffraction imaging,³⁹ and fringe projection profilometry.⁴⁰

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (91850202, 92150301, 12074121, 62105101, 62175066, 11727810, 12034008, 12274129, 12274139); and Science and Technology Commission of Shanghai Municipality (21XD1400900, 20ZR1417100, 21JM0010700).

References

1. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
2. S. Manley et al., "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nat. Methods* **5**(2), 155–157 (2008).
3. M. J. Rust, M. Bates, and X. W. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–795 (2006).
4. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
5. T. Dertinger et al., "Superresolution optical fluctuation imaging with organic dyes," *Angew. Chem. Int. Ed.* **49**(49), 9441–9443 (2010).
6. M. Weber et al., "MINSTED fluorescence localization and nanoscopy," *Nat. Photonics* **15**(5), 361–366 (2021).
7. F. Balzarotti et al., "Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes," *Science* **355**(6325), 606–612 (2017).
8. H. Zhang, M. Zhao, and L. L. Peng, "Nonlinear structured illumination microscopy by surface plasmon enhanced stimulated emission depletion," *Opt. Express* **19**(24), 24783–24794 (2011).
9. L. Reymond et al., "SIMPLE: structured illumination based point localization estimator with enhanced precision," *Opt. Express* **27**(17), 24578–24590 (2019).
10. L. Schermelleh et al., "Super-resolution microscopy demystified," *Nat. Cell Biol.* **21**(1), 72–84 (2019).
11. K. F. Sonnen et al., "3D-structured illumination microscopy provides novel insight into architecture of human centrosomes," *Biol. Open* **1**(10), 965–976 (2012).
12. A. R. Nair et al., "The microcephaly-associated protein Wdr62/CG7337 is required to maintain centrosome asymmetry in *Drosophila* neuroblasts," *Cell Rep.* **14**(5), 1100–1113 (2016).
13. M. A. Ricci et al., "Chromatin fibers are formed by heterogeneous groups of nucleosomes *in vivo*," *Cell* **160**(6), 1145–1158 (2015).
14. A. N. Boettiger et al., "Super-resolution imaging reveals distinct chromatin folding for different epigenetic states," *Nature* **529**(7586), 418–422 (2016).
15. C. A. Wurm et al., "Nanoscale distribution of mitochondrial import receptor Tom20 is adjusted to cellular conditions and exhibits an inner-cellular gradient," *Proc. Natl. Acad. Sci. U. S. A.* **108**(33), 13546–13551 (2011).
16. H. D. Wang et al., "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nat. Methods* **16**(1), 103–110 (2019).
17. R. Chen et al., "Deep-learning super-resolution microscopy reveals nanometer-scale intracellular dynamics at the millisecond temporal resolution," *bioRxiv* 2021.10.08.463746 (2021).
18. C. Qiao et al., "Evaluation and development of deep neural networks for image super-resolution in optical microscopy," *Nat. Methods* **18**(2), 194–202 (2021).
19. P. Llull et al., "Coded aperture compressive temporal imaging," *Opt. Express* **21**(9), 10526–10545 (2013).
20. M. Qiao, X. Liu, and X. Yuan, "Snapshot temporal compressive microscopy using an iterative algorithm with untrained neural networks," *Opt. Lett.* **46**(8), 1888–1891 (2021).
21. A. Paliwal and N. K. Kalantari, "Deep slow motion video reconstruction with hybrid imaging system," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(7), 1557–1569 (2020).
22. H. Z. Jiang et al., "Super SloMo: high quality estimation of multiple intermediate frames for video interpolation," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 9000–9008 (2018).
23. T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM J. Imaging Sci.* **2**(2), 323–343 (2009).

24. P. Tseng and S. Yun, "A coordinate gradient descent method for nonsmooth separable minimization," *Math. Prog.* **117**(1–2), 387–423 (2009).
25. K. Zhang, W. M. Zuo, and L. Zhang, "FFDNet: toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.* **27**(9), 4608–4622 (2018).
26. M. Tassano et al., "FastDVDnet: towards real-time deep video denoising without flow estimation," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 1354–1363 (2020).
27. J. H. Lu and M. L. Liou, "A simple and efficient search algorithm for block-matching motion estimation," *IEEE Trans. Circuits Syst. Video Technol.* **7**(2), 429–433 (1997).
28. J.-Y. Bouguet, "Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm," *Intel Corp.* **5**(4), 1–10 (2001).
29. G. R. Wang, F. Yang, and W. Zhao, "There can be turbulence in microfluidics at low Reynolds number," *Lab Chip* **14**(8), 1452–1458 (2014).
30. E. E. Michaelides, "Brownian movement and thermophoresis of nanoparticles in liquids," *Int. J. Heat Mass Transf.* **81**, 179–187 (2015).
31. F. Z. Zhuang et al., "A comprehensive survey on transfer learning," *Proc. IEEE* **109**(1), 43–76 (2021).
32. A. Creswell et al., "Generative adversarial networks an overview," *IEEE Signal Process. Mag.* **35**(1), 53–65 (2018).
33. D. Ulyanov et al., "Deep image prior," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 9446–9454 (2018).
34. R. J. Yang, L. M. Fu, and H. H. Hou, "Review and perspectives on microfluidic flow cytometers," *Sens. Actuator B Chem.* **266**, 26–45 (2018).
35. M. Schrader et al., "The different facets of organelle interplay—an overview of organelle interactions," *Front. Cell. Dev. Biol.* **3**, 56 (2015).
36. R. D. Vale, "The molecular motor toolbox for intracellular transport," *Cell* **112**(4), 467–480 (2003).
37. Y. Y. Gong et al., "High-speed recording of neural spikes in awake mice and flies with a fluorescent voltage sensor," *Science* **350**(6266), 1361–1366 (2015).
38. P. Schelkens et al., "Compression strategies for digital holograms in biomedical and multimedia applications," *Light Adv. Manuf.* **3**, 40 (2022).
39. Z. Y. Chen et al., "Physics-driven deep learning enables temporal compressive coherent diffraction imaging," *Optica* **9**(6), 677–680 (2022).
40. Y. Hu et al., "Microscopic fringe projection profilometry: a review," *Opt. Lasers Eng.* **135**, 106192 (2020).

Yilin He is a PhD student at State Key Laboratory of Precision Spectroscopy, East China Normal University under the supervisions of Prof. Shian Zhang. His research focuses on high-speed super-resolution microscopy.

Yunhua Yao is an associate professor at State Key Laboratory of Precision Spectroscopy, East China Normal University (ECNU). He received his PhD in optics from ECNU in 2018. His current research interest focuses on high-speed super-resolution microscopy and ultrafast optical imaging.

Dalong Qi is a young professor at State Key Laboratory of Precision Spectroscopy, East China Normal University (ECNU). He received his PhD in optics from ECNU in 2017. His current research interest focuses on ultrafast optical and electronic imaging techniques and their applications.

Yu He is a PhD student at State Key Laboratory of Precision Spectroscopy, East China Normal University under the supervisions of Prof. Shian Zhang. His research focuses on high-speed super-resolution microscopy.

Zhengqi Huang is a PhD student at State Key Laboratory of Precision Spectroscopy, East China Normal University under the supervisions of Prof. Shian Zhang. His research focuses on high-speed super-resolution microscopy.

Pengpeng Ding is a PhD student at State Key Laboratory of Precision Spectroscopy, East China Normal University under the supervisions of Prof. Shian Zhang. His research focuses on ultrafast optical imaging.

Chengzhi Jin is a PhD student at State Key Laboratory of Precision Spectroscopy, East China Normal University under the supervisions of Prof. Shian Zhang. His research focuses on ultrafast optical imaging.

Chonglei Zhang is a professor from Nanophotonics Research Center, Shenzhen University. He received his PhD in physics from Nankai University in 2007. His current research interest focuses on super-resolution microscopy and surface plasmon resonance.

Lianzhong Deng is an associate professor from State Key Laboratory of Precision Spectroscopy, East China Normal University (ECNU). He received his PhD in optics from ECNU in 2008. His current research interest focuses on ultrafast optical and electronic imaging techniques and their applications.

Kebin Shi is a professor from State Key Laboratory for Mesoscopic Physics, Peking University. He received his PhD from Pennsylvania State University. His current research interest focuses on nonlinear photonics and biomedical imaging.

Zhenrong Sun is a professor of the State Key Laboratory of Precision Spectroscopy, East China Normal University (ECNU). He received his PhD in physics from ECNU in 2007. His current research interest focuses on ultrafast dynamics of clusters and ultrafast optical imaging.

Xiaocong Yuan is a professor and the director of Nanophotonics Research Center, Shenzhen University. He received his PhD in physics from King's College London in 1994. His current research interest focuses on optical manipulation, high-sensitivity sensor, super-resolution microscopy, and surface-enhanced Raman spectroscopy.

Shian Zhang is a professor and the deputy director of State Key Laboratory of Precision Spectroscopy, East China Normal University (ECNU). He received his PhD in optics from ECNU in 2006. His current research interest focuses on ultrafast optical imaging, high-speed super-resolution microscopy, and light field manipulation.