Supplementary Information for

Full-color 3D Visualization with Janus metafiber

Haotian Xu[1†], Tianyue Li[2†], Haotian Ding[1], Shaochen Fang[1], Xiangyu Ma[1], Haojie Chen[3], Ye Chen[3], Yifeng Xiong[1], Shuming Wang[1*], Fei Xu[1,4,5*]

1. National Laboratory of Solid State Microstructures, College of Engineering and Applied Sciences, School of Physics, Nanjing University, Nanjing, 210023, China

2. Department of Physics and State Key Laboratory of Optical Quantum Materials, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, 999077, China

3. MIIT Key Laboratory of Aerospace Information Materials and Physics, State Key Laboratory of Mechanics and Control for Aerospace Structures, College of Physics, Nanjing University of Aeronautics and Astro-nautics, Nanjing 211106, China

4. Hefei National Laboratory, Hefei 230088, China

5. Shenzhen Research Institute of Nanjing University, Shenzhen, 581000, China

[†]These authors are contributed equally: H. Xu, T. Li

[*]Correspondence: wangshuming@nju.edu.cn; feixu@nju.edu.cn

This PDF file includes:

Supplementary Note 1 to 10
Supplementary Figures 1 to 12
Supplementary References

Supplementary Note 1. The optical configuration of imaging system

The optical configuration diagram employed in the actual testing is presented in **Fig. S1**. We utilize a bidirectional optical path configuration to achieve dual-mode imaging. Three-dimensional depth imaging is realized by displacing the focused spot on the fiber input end-face via a spatial light modulator (SLM). Subsequently, at the output end, a secondary phase metasurface amplifies the angle and collimates the output beam. This is then collected by multimode fiber and directed into an avalanche photodetector (APD). When switching to two-dimensional full-color imaging mode, we merely require the white light LED to be collimated and directed into BS_2. This couples the illumination light into the multimode fiber to illuminate the imaging object. The imaging fiber collects the color image information, and the output image passes through BS_1 to enter the CCD for two-dimensional imaging. During actual testing, no optical path components or configuration changes are required. Rapid switching between three-dimensional and two-dimensional imaging functions is achieved solely through electrically controlled, time-alternating activation of the LED and SLM corresponding to the phase map. The current acquisition rate is constrained by the repetition rate of the pulsed laser; future implementation with a suitable laser will enable video-grade data acquisition rates.

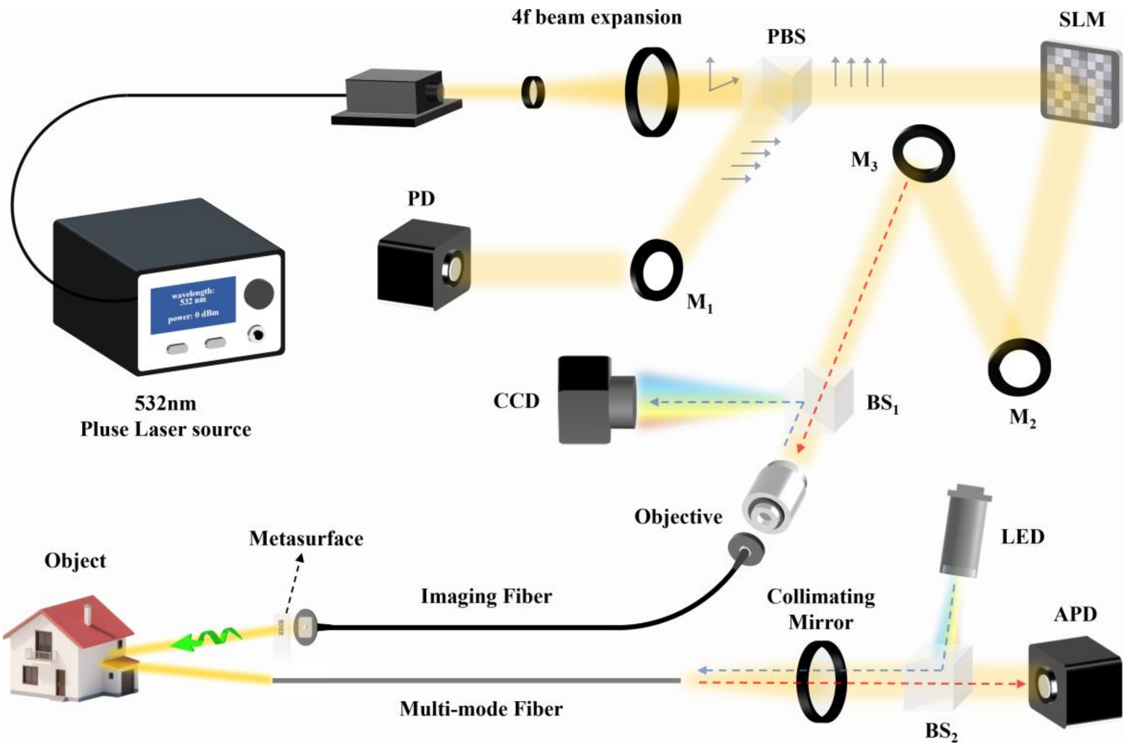

Figure S1. Optical configuration of FCSV. The meanings of the symbols are as follows: **PBS**: polarized beam splitter. **SLM**: spatial light modulator. **M_1**: reflector 1. **M_2**: reflector 2. **M_3**: reflector 3. **PD**: photodetector. **BS_1**: beam splitter 1. **CCD**: charge-coupled device. **BS_2**: beam splitter 2. **APD**: avalanche photodetector. **LED**: light emitting diode.

Supplementary Note 2. The optical configuration of imaging system

In this work, we employed a bidirectional switching dual-function multiplexing system configuration. By utilizing multimode fiber to transmit white light and collect pulsed laser signals, we achieved two-dimensional full-color imaging and three-dimensional depth imaging respectively. The choice of multimode fiber for illumination was driven by considerations of overall system integration, as well as constraints on endoscope dimensions in potential future endoscopic applications. Concurrently, to collect pulse signals with sufficient APD-detectable intensity in the time-of-flight imaging configuration, the core diameter of the multimode fiber cannot be reduced indefinitely. This imposes a certain constraint on further miniaturization of the system.

The front-end configuration of the actual fiber imaging apparatus is illustrated in **Fig. S2**. The imaging fiber is positioned approximately 600 μm from the secondary phase metasurface, with a spacing of roughly 1 cm between the imaging fiber and the multimode fiber. However, the primary constraint on the distance between them is the size of the fixture, chiefly comprising the displacement stage fixtures securing the imaging fiber and the metasurface sample. These mechanical structures obstruct the placement of the multimode fiber, thus preventing the two fibers from being positioned in close proximity. Should more compact fixtures be employed, or should the fibers and metasurface samples be fully integrated in subsequent practical applications, the distance between them could be reduced to the sum of their radii—approximately 1.5 mm—enabling imaging demonstrations within most confined spaces.

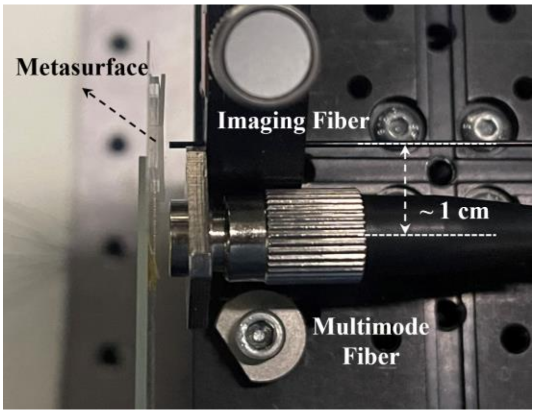

Figure S2. The diagram of the actual FCSV device.

Metasurface fabrication process

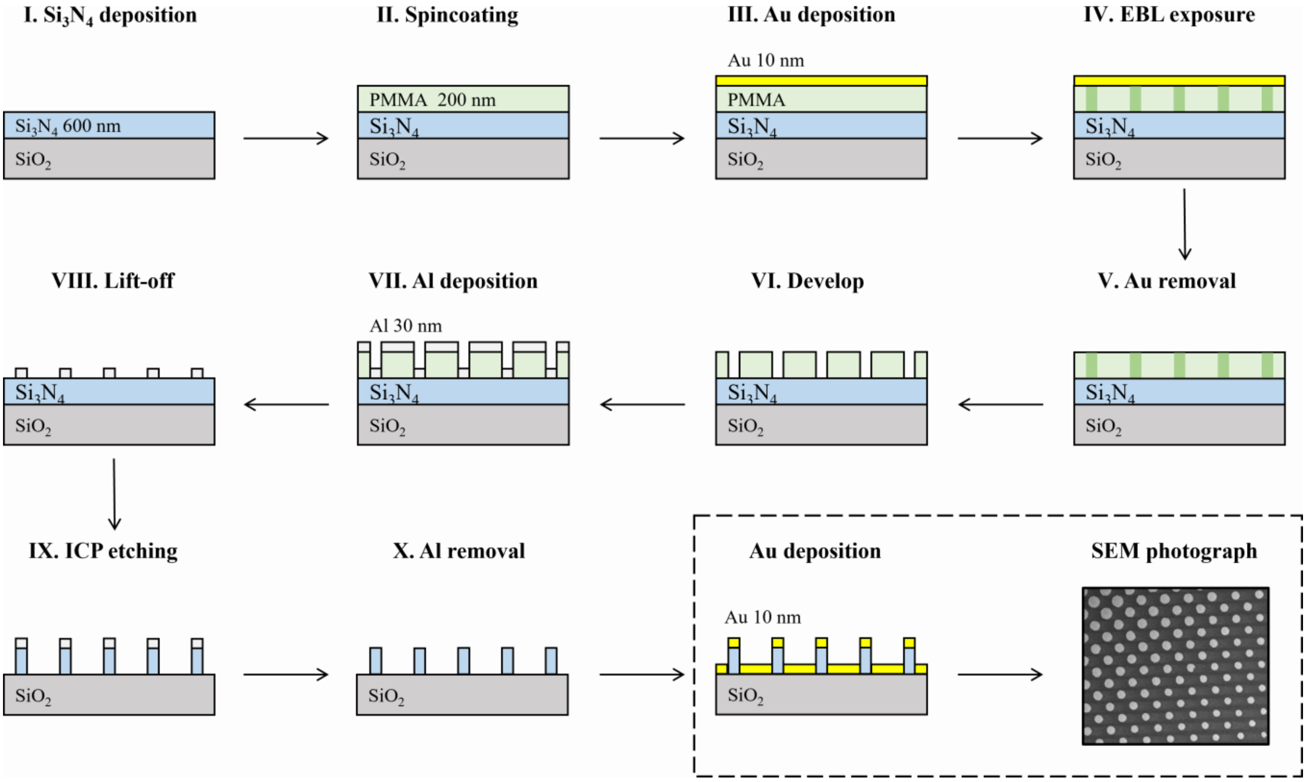

Figure S3. Fabrication process of metasurface. We used a standard CMOS-compatible fabrication process to prepare the metasurface with a sequence of Si_3N_4 deposition, spin-coating, Au deposition, EBL exposure, Au removal, developing, Al deposition, lift-off, ICP etching, Al removal and final SEM photographing. The first ten steps are the original fabrication process and the steps in the dashed box are the additional SEM photographing process.

Multi-wavelength PSF results within 0° to 30°

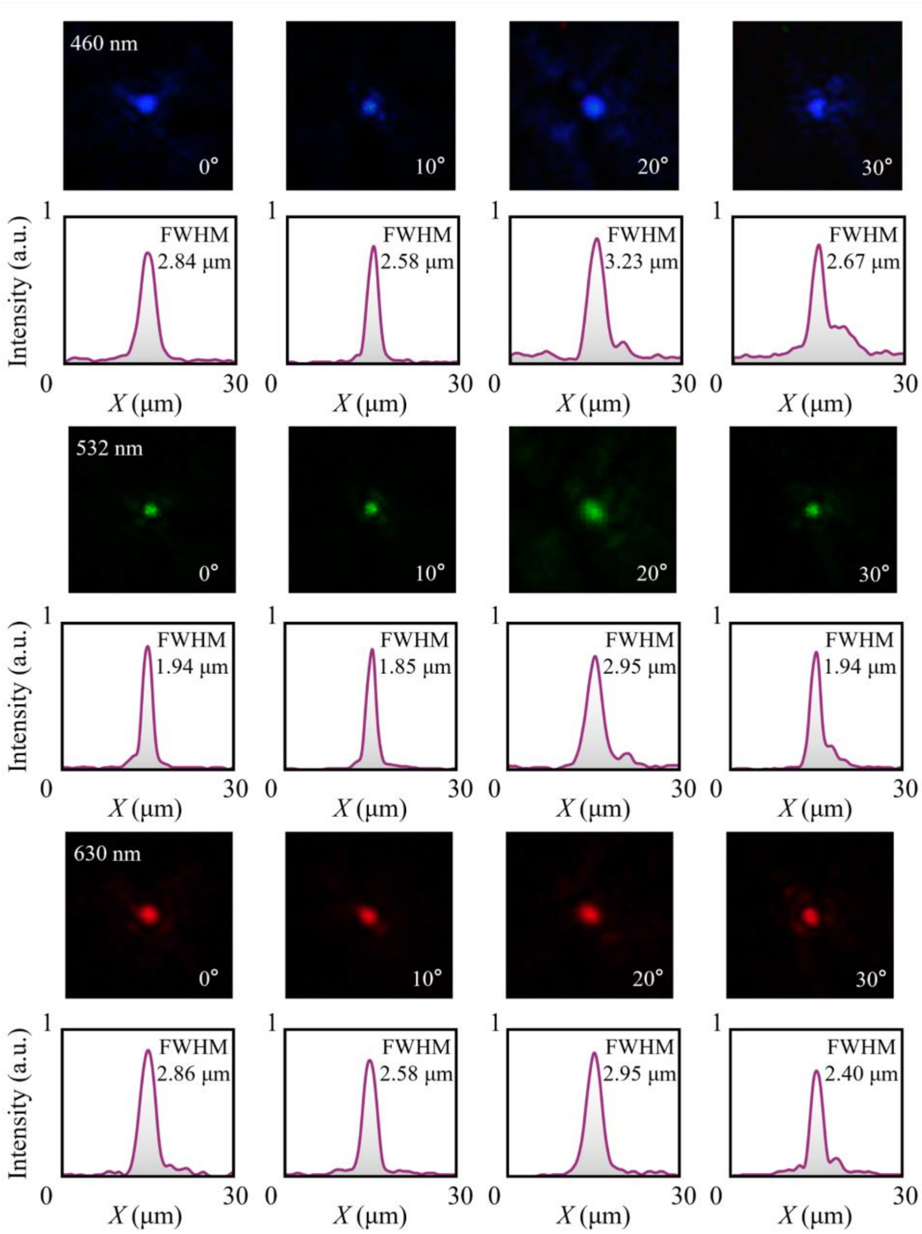

Figure S4. The PSF result and corresponding FWHM value of three wavelengths. The PSF response results and FWHM values in 460 nm, 532 nm, and 630 nm are displayed from top to bottom, respectively. From the 2D image distribution we can found that the PSF response are consistent within the incident angle from 0° to 30°. The specific one-dimensional intensity profile extracted from the PSF result are shown in the below.

Supplementary Note 3. The image of the exit spots on the end-face

An important characteristic of a quadratic phase metasurface is that its incident focal point shifts with changes in the incident angle. Therefore, we first verified this characteristic. Based on the principle of optical path reversibility, we first determined the SLM phase map corresponding to the target angle, and then used a CCD to capture the exit beam patterns at the end-face of the fiber bundle under these corresponding phases, as shown in **Figure S5**. The left side shows the bright-field image of the end-face illuminated by a defocused beam, used to determine the boundary of the fiber bundle. We then captured the spot patterns within a ± 30° range, which shifted with the exit angle. The specific shift values were measured and listed in **Figure 4b** of the main text, which are consistent with the theoretically predicted shift values.

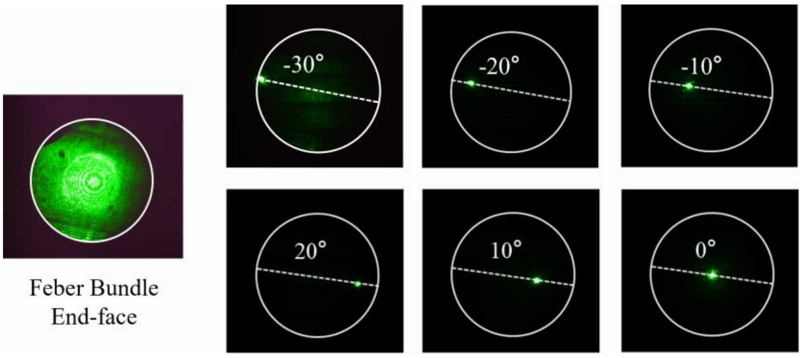

Figure S5. The offset light spot of different angels and the bright-field image of the whole end-face. The exit spots on the end-face between -30° and 20°.

Supplementary Note 4 The effective FOV analysis of system in 3D depth imaging

The theoretical limiting field of view (FOV) for this secondary phase metasurface is 60°, a characteristic fully realized in the imaging field of view during 2D imaging. However, when performing 3D depth imaging, considerations regarding reflection intensity and the large-angle light collection efficiency of multimode fiber necessitate balancing the power intensity receivable by the APD during scanning imaging tests to determine the practical FOV. The multi-angle test results for the collimated output beam power of this imaging system are presented in **Fig. S6** Due to aperture clipping at large angles and background light attenuation, the beam power projected onto the object decreases to approximately half its 0° value at 20°. At even larger angles, the weaker power received at the APD results in poor resolution. Consequently, we define ±20° as the effective FOV for this imaging system. In the **Fig. 4** in main text, we present test results solely within this range to correspond to the actual field of view under this imaging mode.

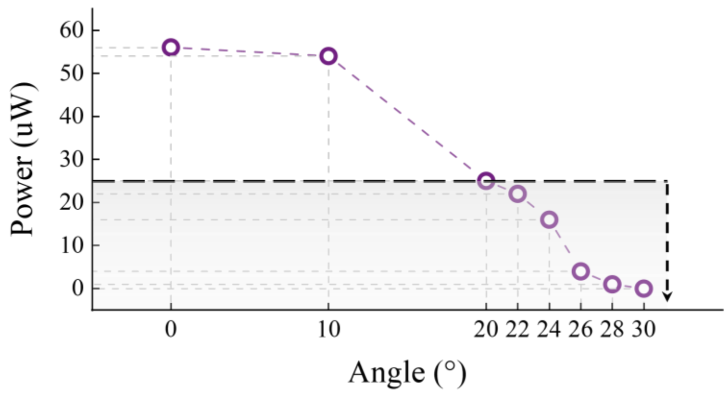

Figure S6. The power testing at different deflection angles.

Supplementary Note 5. The fitting process of output spot

When testing the 3D depth imaging performance of this system, we must determine the divergence angle of the collimated output beam, which dictates both the collimation capability of this secondary phase metasurface and its depth pixel resolution. Typically, the divergence angle α can be calculated simply by measuring the sizes of the output spot and the spot on the object, as illustrated in **Fig. S7a**:

$$\alpha = 2\theta = 2\tan^{-1}\frac{\frac{l_2}{2}-\frac{l_1}{2}}{d} \tag{1}$$

Here, d denotes the object depth, l_2 represents the full width at half maximum (FWHM) of the spot on the object, and l_1 denotes the FWHM of the emitted spot. When the emission is oblique, the projection dimension along the beam's normal direction is substituted for calculation. At the exit light source, the beam diameter is approximated as 600 μm. The spot size on the object light source cannot be directly obtained, as it has been sampled and segmented by the fibre core (**Fig. S7b**). Therefore, a two-dimensional Gaussian fit is required on the sampled spot to derive its original Gaussian spot full width at half maximum (FWHM), thereby calculating its divergence angle. The complete fitting process is illustrated below:

- Identify local extrema points in the post-sampling light spot as the dataset for subsequent Gaussian spot restoration.
- Compute initial parameters from this dataset and fit a two-dimensional Gaussian function using non-linear least squares.
- Reconstruct the light spot from the fitted data and estimate its full width at half maximum.

The detailed process is illustrated in **Fig. S7c**. In the **Fig. 4d** in the main text, the green image on the left represents the speckle image captured by the CCD on the object, while the Gaussian spot on the right is the reconstruct image after Gaussian fitting.

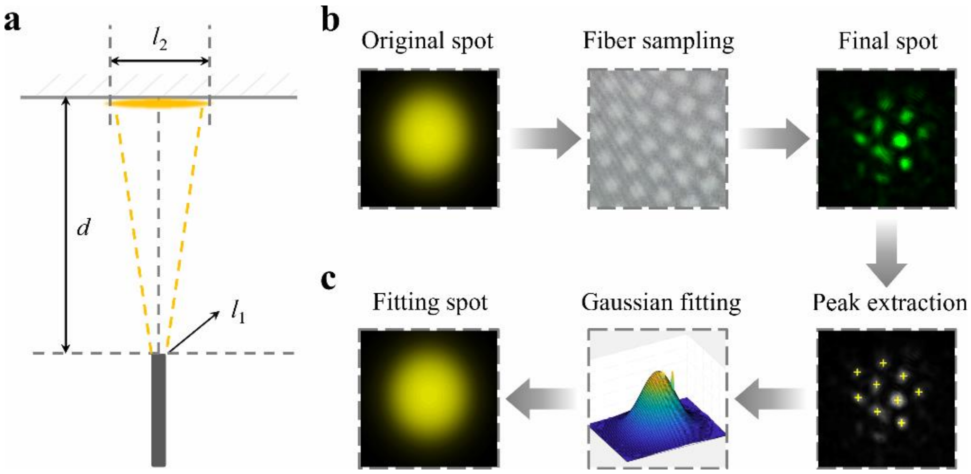

Figure S7. The 2D Gaussian fitting process of output spot. (a) Schematic diagram of the emitted light spot diffusion propagation. **(b)** The original light spot is sampled by the imaging fiber cores and emitted. **(c)** Perform multi-point two-dimensional Gaussian fitting on the emitted light spot to calculate the beam divergence angle.

Supplementary Note 6. The analysis of focusing efficiency in different angles

In practical scanning imaging performance testing, we calculate the focusing efficiency by dividing the spot power P_f focused onto the object by the power P_t transmitted through the secondary phase metasurface, namely:

$$E_f = \frac{P_f}{P_t} \tag{2}$$

At the output end, the power contained within P_t comprises two components: the **collimated beam power** representing the focused spot, and the **background light power** representing the background spot. The focused spot is formed by the coupling of light directly focused onto the fiber input end-face, while the background light originates from the light field spilled from the focused spot into the surrounding area, combined with various reflected light at the front-end focusing surface and stray light within the optical path. After diffusing over a certain distance, these two light components will form the distribution shown in **Fig. S8a**, where the brightness of the focused spot is significantly higher than that of the background spot. During actual power testing, considering the geometric dimensions of the spot, a spatially adaptive visible light powermeter must be employed (**Fig. S8b**). The measured power P_f thus encompasses both the focused spot power and a portion of the transmitted background light power:

$$P_f = P_{focus} + P_{background} \tag{3}$$

During multi-angle testing, we assume that the power P_{focus} of the collimated beam propagating through free space incurs no significant attenuation. Its power distribution evolves strictly in accordance with its divergence angle, **remaining constant throughout the calculations**. However, the background light component maintains a largely consistent total power, thus remaining unchanged in the denominator of the calculation. Yet, the portion propagated through space, denoted as $P_{background}$, is influenced by factors such as the level of power coupling between fibers and crosstalk from multimode spots within the fiber. Consequently, its behaviour varies across different angles and exhibits a degree of randomness, with its evolution determined by the transmission process within the imaging fiber. Subsequent testing repeatedly measured background spot power at various angles and positions, with specific values $P_{background}$ fluctuating between 1 to 6 μW. This introduces a degree of error into overall efficiency testing. During testing, the output power P_t at the fiber output end, achieved based on certain optical path attenuation, was approximately 70 μW, while the focused spot power was approximately 40 μW. The fluctuation component of the background spot power, $\Delta P_{background}$, was taken as 5 μW. At this point, the difference in focusing efficiency ΔE_f measured at different angles was:

$$\Delta E_f = \frac{P_{fmax} - P_{fmin}}{P_t} = \frac{\Delta P_{background}}{P_t} = \frac{5\ uW}{70uW} = 7.2\% \tag{4}$$

This closely aligns with the maximum difference in focusing efficiency observed at different angles in **Fig. 4f** in the main text. Consequently, **the variation in focusing efficiency with angle exhibits not strict symmetry centred at 0°, but rather a degree of fluctuation.** This behaviour stems particularly from the computational methodology employed and the propagation characteristics of the light field. In contrast, the transmission efficiency and overall system efficiency are governed by the off-axis attenuation law of the entire optical system, thereby exhibiting a distinctly centred symmetric distribution.

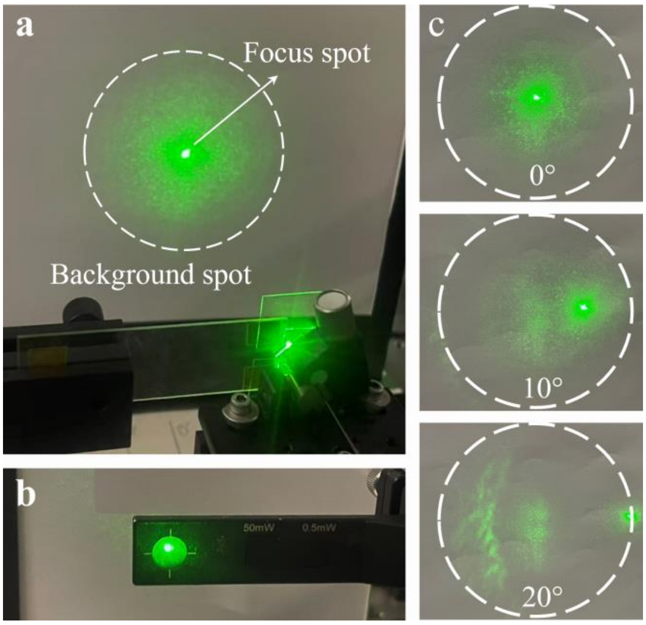

Figure S8. The spot component diagram of FCSV system. (a) Actual image of the focused light spot and background light. **(b)** Testing the power of the focused spot with a power meter. **(c)** Focus spot and background light distribution at different angles.

Supplementary Note 7. The sine interpolation of original data

Figure S9a shows the single-pulse data of the original acquired signal, with an acquisition rate of 1 GHz and a corresponding time resolution of 1 ns. We used a sine interpolation method to interpolate and expand the data points to improve the time-domain resolution. Sine interpolation is a method that uses sine trigonometric functions for data interpolation. Due to its smoothness and periodicity, it can construct a reasonable curve between known and unknown points. Essentially, it interpolates discrete signals using the time-domain curve of a low-pass filter. Specifically, in the time domain, it manifests as the convolution of the original signal with the low-pass filter function h[k]:

$$y[t] = x[t] * h[t] \tag{5}$$

Where $x[t]$ is the function to be interpolated, $h[t]$ is the time domain response function of the low-pass filter, and $y[t]$ is the interpolated result. When substituting into the convolution formula and selecting N interpolation points within a certain range, the formula can be simplified as follows:

$$y[t] = \sum_{n=0}^{N-1} x[nT] \frac{\sin\left(\frac{\pi}{T}(t - nT)\right)}{\frac{\pi}{T}(t - nT)} \tag{6}$$

Where T is the sampling interval. Using this method, we performed interpolation calculations on a single pulse with a discrete point count $n = 32$ and a sampling interval $T = 1$ ns, with the results shown in **Figure S9b**. As can be seen, the interpolated curve is relatively smooth, and the trend is almost identical to the original data. Therefore, we used the interpolated pulse samples for flight time difference calculations. We use the point coordinates at half of the maximum value in the rising edge as the pulse coordinates for comparing the signal with the reference pulse. During calculation, the first half of the pulse peak is subtracted by half the maximum value and its absolute value is taken. The coordinate value is then determined through minimum value search, and the processed pulse image is shown in **Figure S9c**.

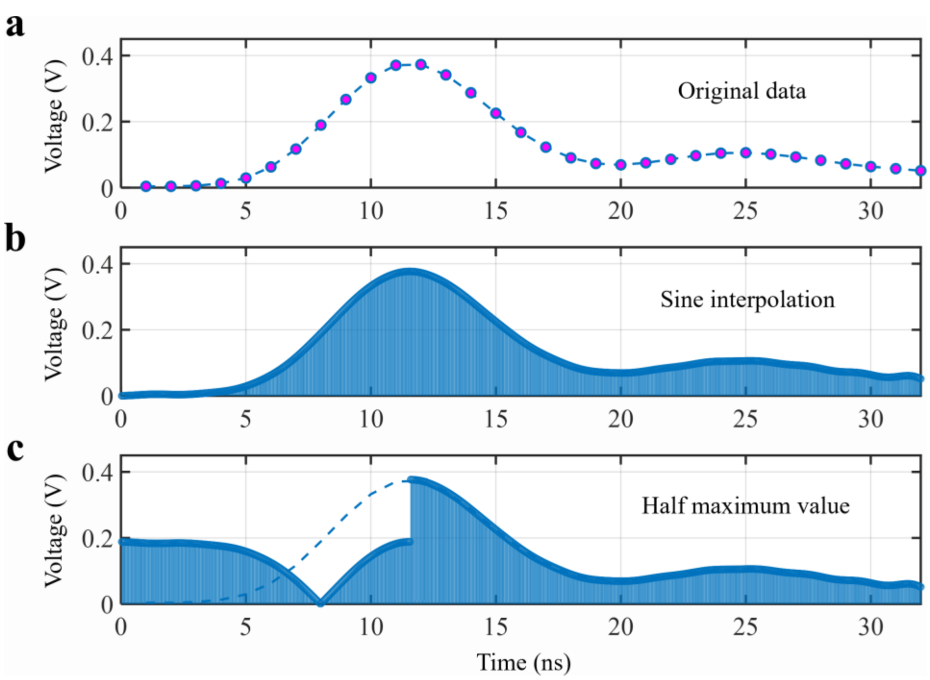

Figure S9. The comparison between the (a) original data, (b) sine interpolation data and (c) processed data.

Supplementary Note 8. The processing of original data

In a standard TOF system, depth resolution is directly determined by the bandwidth of the detector, and higher sampling frequency will result in higher temporal resolution. Specifically, in our work, the sampling frequency of the APD used is 1 GHz (corresponding to a sampling time interval Δt of 1 ns). Assuming the scene depth is d, the minimum distinguishable depth difference is Δd, the speed of light is c, the distance between the imaging fiber and the multimode fiber is h, and the offset distance of the pulse hitting position on the object relative to the center point is l (i.e., oblique emission).

When the imaging fiber is closely adjacent to the receiving multi-mode fiber, as shown in **Fig. S10a**, there are:

$$\Delta t \cdot c = 2\Delta d \tag{7}$$

$$\Delta d = \frac{\Delta t \cdot c}{2} = \frac{1\times10^{-9}s\times3\times10^{10}cm\cdot s^{-1}}{2} \approx 15\ cm \tag{8}$$

In the case of adjacent and normal emission, the minimum distinguishable depth accuracy is 15 cm.

In practical experiments, due to the limitations of mechanical components, there is a certain distance between two fibers, as shown in **Fig. S10b**. At this time, there are:

$$\Delta t \cdot c = \Delta d + \sqrt{(\Delta d + d)^2 + h^2} - \sqrt{d^2 + h^2} \tag{9}$$

Taking the interval h as 1 cm, the calculation yields $\Delta d = 15.015$ cm, which is almost the same as when there is no interval. At the same time, we also considered the situation when emitting obliquely, as shown in **Fig. S10c**. The maximum offset value lmax of the emission point relative to the center point is taken as 2.5 cm. At this time, the relevant formula becomes:

$$\Delta t \cdot c = \sqrt{(\Delta d + d)^2 + l^2} + \sqrt{(\Delta d + d)^2 + (l + h)^2} - \sqrt{d^2 + h^2} - \sqrt{d^2 + (l + h)^2} \tag{10}$$

After calculation, $\Delta d = 15.27$ cm can be obtained, which is basically consistent with the simplified case. Therefore, we approximately assume that the intrinsic depth resolution of the system is about 15 cm.

In this work, we used sine interpolation method to improve depth resolution, as shown in **Supplementary Note 7**. We used 32-point interpolation between the original data points, so theoretically the depth resolution of the system should be improved to $\Delta d/32 = 15/32 \approx 0.5$ cm.

However, excessive interpolation may lead to point selection errors due to existing system and measurement errors, resulting in a decrease in resolution. In actual testing, we found that the final verified reliable depth resolution is about 2 cm, which is approximately 1/8 of the system's intrinsic depth resolution.

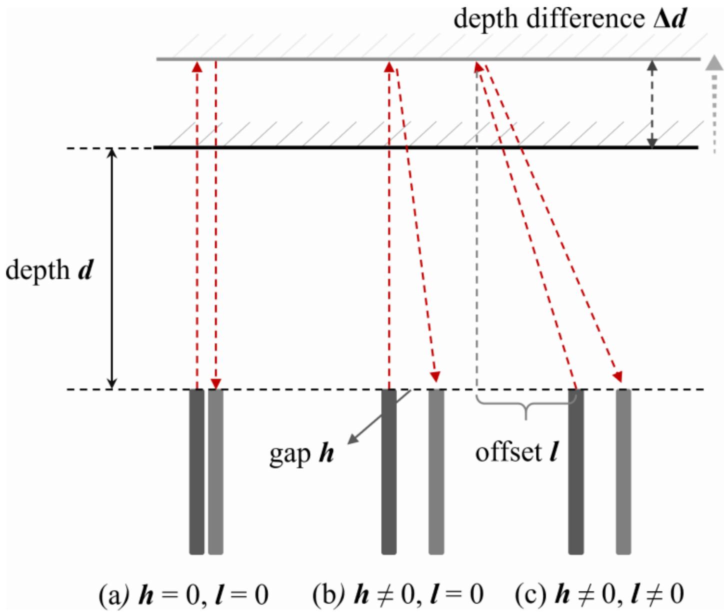

Figure S10. Schematic diagram of depth information collection method. (a) Adjacent and normal emission situation. (b) Interval and normal emission situation. (c) Interval and oblique emission situation.

Supplementary Note 9. The proc4essing of original data

Considering the FWHM of the output light spot, we used a 50×50 pixel array scan field, and the original data obtained is shown on the far left of **Figure S11**. A bilateral filtering algorithm was used to denoise and smooth the image, removing random errors and noise effects from the measurement. We then used the Imresize function in Matlab to upscale the image, increasing the pixel resolution to 200×200 to improve image quality and ensure resolution compatibility for subsequent synthetic processing with full-color images. Finally, we applied a bilateral filtering process to the image, yielding the results presented in **Figure 5b** in the main text. As shown in the figure, although we performed up-sampling and smoothing filtering on the image, these processes merely enhanced the image smoothness and sampling resolution based on the original image, without altering or artificially modifying the image details reflected in the original data.

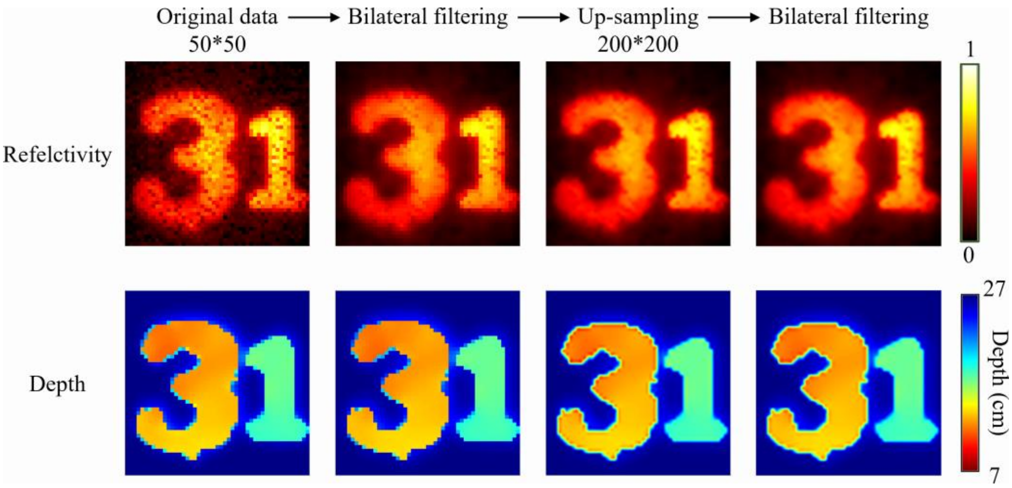

Figure S11. The comparison between the original data, bilateral filtering data and up-sampling data.

Supplementary Note 10. The integration strategy of metafiber system

For practical applications—particularly in endoscopic imaging—metasurface samples must be robustly integrated with optical fibers to ensure mechanical stability and operational reliability.

In the metafiber system, the most ideal integration solution is to directly perform metasurface fabrication on the fiber end-face, which can accurately control the spacing and position relationship between them and ensure their integrity and stability. However, the current mainstream micro-nano processing equipment are most designed for flat samples, so it is difficult to be compatible with fiber devices with large aspect ratio characteristics. A more convenient alternative is to first process the metasurface sample on a flat substrate, followed by assembly and packaging with optical fibers. This method can ensure both the fabrication accuracy of the metasurface and the feasibility of the integrated process, as shown in **Figure S12a**.

The fabricated metasurface sample is diced via femtosecond laser ablation, followed by alignment and integration with an optical fiber. Two approaches are available to realize the interconnection between these components. The first strategy involves direct bonding to the fiber end-face, or even the surface of fiber jumpers with ultraviolet-curable adhesive severing as the transition layer. This adhesive features a refractive index closely matched to that of the silicon dioxide substrate. We have previously validated this approach in our earlier work[1], as illustrated in **Figure S12b**. The alternative method employs a quartz capillary tube as a sleeve to fix their relative positions. In this case, the diced metasurface sample must first be secured to the quartz capillary, after which the optical fiber is inserted into the sleeve and fixed at an appropriate distance from the metasurface. This scheme is better suited for the integration of bare fibers and samples, as direct alignment struggles to prevent fiber bending or tilting, thereby introducing positional errors.

We present the integrated sample images of imaging fiber and planar sample using this universal method in **Figure S12c** and **d**, where **Figure S12c** shows the integration of the lens group and imaging fiber through a capillary tube, and **Figure S12d** shows the image of integration sample of the secondary phase metasurface in this paper. In **Figure S12d**, three distinct structural regions are clearly discernible: bare fiber areas, sleeve areas, and the front planar sample layer. The overall diameter of the sample is about 1.2 mm, and the rigid length (i.e. sleeve length + substrate thickness) is about 3 mm. Through process optimization, its rigid length can be further minimized to enhance both system-level integration density and mechanical adaptability in highly tortuous endoscopic configurations.

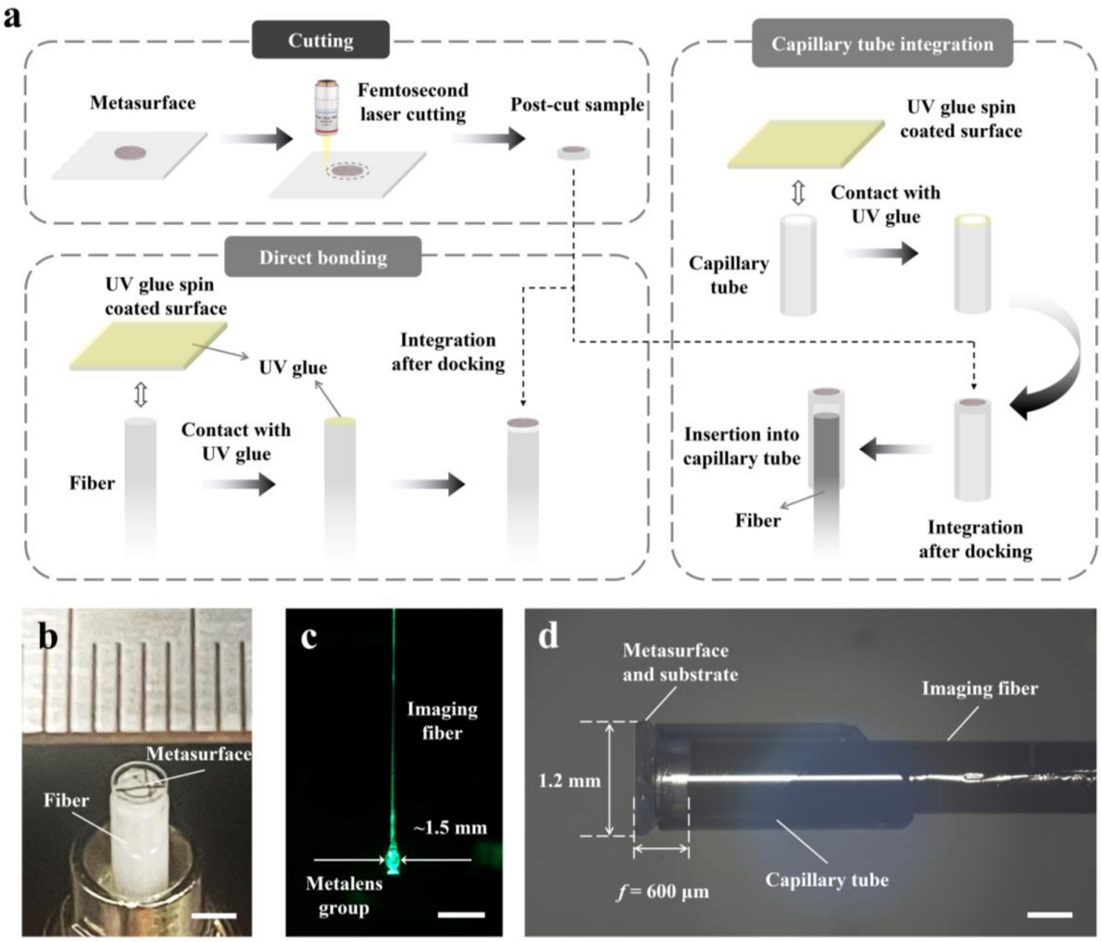

Figure S12. The integration strategy of metafiber system and related results. (a) The diagram of the integration process between metasurface and optical fiber. There are two main pathways: UV-curable adhesive assisted direct bonding and capillary tube–assisted alignment and packaging. **(b)** The image of integration sample between single metasurface and fiber jumper using direct bonding method. Scale bar, 2 mm. **(c)** The image of integration sample between metalens group and imaging fiber using capillary tube integration method. Scale bar, 5 mm. **(d)** The image of integration sample between secondary phase metasurface and imaging fiber using capillary tube integration method. Scale bar, 500 μm.

Reference

1. Xu H, Li T, Gao G, Fang S, Liu H, Ding H, Chen X, Xiong Y, Wang S, Xu F. Metasurface-Integrated Pattern-Preserved Fiber Mode Separator. *Advanced Functional Materials* **35**, 2505081 (2025).