Peer Review File

Full-color 3D Visualization with Janus Metafiber

Corresponding Author: Professor Fei Xu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript reports the design, fabrication, and characterisation of a dual function fibre optic imaging system capable of delivering wide angle 2D imaging in the visible range via a metalens integrated multicore fibre bundle, as well as 3D depth imaging through a multimode fibre using a time of flight approach. The work is primarily engineering driven, enriching fibre optic imaging capabilities toward multifunctional imaging fibres. The design and results are interesting and technically sound. I support its publication in Nature Communications; however, the authors should address the following concerns and suggestions before acceptance.
1. In Figure 1 (or elsewhere), an image of the actual fibre device should be provided, and the lateral separation between the fibre bundle and the multimode fibre should be specified.
2. The manuscript should tone down its claims regarding endoscopy, as no endoscopic imaging experiments are demonstrated. A title such as “Full colour stereoscopic Janus metafibre” would better reflect the current scope. Applying this metafibre to endoscopic imaging is a separate endeavour and would require biomedical imaging results.
3. In the Abstract, the sentence beginning with “In this work, we propose …” is overly long. The term “multiplexing” is also used in a confusing way. The manuscript does not demonstrate “forward and reverse multiplexing” or “time division multiplexing.” These buzzwords should be removed, and the methods should be described using clearer and more accurate terminology to improve accessibility.
4. The SEM image in Figure 2f shows connected meta atoms. Will this affect the PSF or overall imaging performance? Please clarify.
5. In Figure 3d, why does the PSF FWHM become smaller at a 10 degree incidence angle?
6. In Figure 4f, why is the focusing efficiency higher at –10 degrees, and why is the response not symmetric around normal incidence?
7. The purpose of the Gaussian fitting in Figures 4d and 4e is unclear. How does the green image on the left correspond to the Gaussian spot shown on the right in Figure 4d?
8. The necessity and relevance of edge detection for fibre optic imaging applications should be discussed.
9. The manuscript employs a simple metalens design without chromatic aberration correction, resulting in a smaller PSF at the design wavelength of 532 nm. The authors should clarify whether such a metalens is adequate for full colour imaging. The following paper should be cited in this context:
https://www.nature.com/articles/s41467-022-31902-3

Reviewer #2

(Remarks to the Author)
The manuscript presents an interesting metasurface-based fiber imaging system with promising experimental results. However, several key aspects of the system and its claimed performance require further clarification. The authors are advised to conduct a major revision before the manuscript is reconsidered.

1. Full-color 2D imaging and 3D ToF sensing rely on different illumination sources and detectors (LED/CCD versus laser/SLM/APD). Could the authors clarify whether both modalities can be realized through mode switching within a fixed system configuration, or whether changes in the optical layout or component positioning are required?

2. The manuscript introduces an additional multimode fiber for ToF illumination and signal collection. Could the authors

elaborate on whether this MMF plays an essential role beyond light delivery and collection, and whether it offers clear advantages for system integration or miniaturization given the current architecture?

3. As the metasurface appears to be optimized for a single working wavelength, could the authors clarify how broadband chromatic aberrations influence full-color imaging performance? In particular, how do the wavelength-dependent PSF variations observed in Fig. S2 affect image fidelity, and might dispersion-engineered metasurface designs help improve broadband focusing stability?

4. The Introduction reports a field of view of 60°, whereas the main text focuses on performance within ±20° due to image degradation at larger angles. Could the authors clarify the distinction between the maximum geometrical FOV and the effective FOV with acceptable image quality?

5. The reported improvement in depth resolution relies in part on interpolation techniques. Could the authors provide additional quantitative analysis to clarify how the claimed depth accuracy relates to the intrinsic system resolution?

6. A careful proofreading to address minor typographical and language issues would further improve the clarity of the manuscript.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I appreciate the authors' careful attention to my previous comments. The revised manuscript now includes an optical image of the actual FCSV device, and I am satisfied that all technical issues have been addressed.
However, one critical concern remains: the integration of a metasurface with the fibre system. The current device does not demonstrate a true metafibre, which by definition involves direct integration of a metasurface onto the facet of an optical fibre, a concept well established in the literature. Since the manuscript claims "Full-color 3D Visualization with Janus Metafiber," this integration is essential to justify the central innovation and meet the expectations of Nature Communications.
I strongly encourage the authors to pursue a practical integration strategy that reflects the metafibre concept. This does not necessarily require direct fabrication on the fibre facet; even attaching a metasurface via a thin substrate can serve the purpose, provided the metasurface and fibre function as a unified device.
Without this integration, the technical advancement is significantly diminished, and the current implementation, relying on free-space and fibre-optic components, appears standard and below the threshold for publication in Nature Communications. With integration, however, I believe the manuscript is ready for acceptance.

Reviewer #2

(Remarks to the Author)
The paper has been revised in accordance with the review comments, and this version is recommended for acceptance.

Responses to Reviewer #1's comments

Reviewer #1: This manuscript reports the design, fabrication, and characterization of a dual function fiber optic imaging system capable of delivering wide angle 2D imaging in the visible range via a metalens integrated multicore fiber bundle, as well as 3D depth imaging through a multimode fiber using a time of flight approach. The work is primarily engineering driven, enriching fiber optic imaging capabilities toward multifunctional imaging fibers. The design and results are interesting and technically sound. I support its publication in Nature Communications; however, the authors should address the following concerns and suggestions before acceptance.

Response: We sincerely appreciate the reviewers' insightful comments and constructive suggestions, which have significantly improved the quality of our manuscript. We carefully addressed all the issues and accordingly revised the manuscript with highlight marking for convenience.

1. In Figure 1 (or elsewhere), an image of the actual fiber device should be provided, and the lateral separation between the fiber bundle and the multimode fiber should be specified.

Response: I really appreciate your comment on this problem. The front-end configuration of the actual fiber imaging apparatus is illustrated in **Fig. R1**. 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.

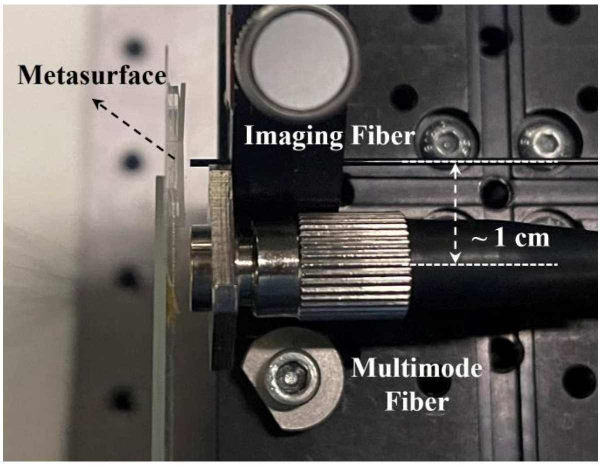

Figure R1. The diagram of the actual FCSV device.

Manuscript revision:

1. We added a sentence in the **Page 5 line 12** of the main text:

The actual picture of the output part of this system is shown in Supporting Information Note 2 and Figure S2).

2. We added Supplementary Note 2, Fig. S2 in Supplementary information.

2. The manuscript should tone down its claims regarding endoscopy, as no endoscopic imaging experiments are demonstrated. A title such as "Full colour stereoscopic Janus metafiber" would better reflect the current scope. Applying this metafiber to endoscopic imaging is a separate endeavour and would require biomedical imaging results.

Response: Thank you for your comments on this question. Considering the common applications of fiber imaging systems, we had broadly termed it "fiber endoscope" in the original main text to correspond with its potential macro-endoscopic applications. We appreciate your correction on this point and acknowledge that defining it as an endoscope without reference to specific endoscopic testing scenarios was indeed inappropriate. Consequently, we have amended the title and the terminology used throughout the main text. The relevant revisions have been highlighted in **RED**.

Manuscript revision:

1. We modified the title of the main text to "Full-color 3D Visualization with Janus metafiber".
2. We modified a sentence in the **Page 1 line 4** of the main text:

Fiber imaging system exhibits high-precision imaging capability in extreme scenarios due to its flexibility and compatibility.

3. We modified a sentence in the **Page 1 line 17** of the main text:

Fiber-optic imaging system brings forth minimally invasive or non-destructive inspection in biomedical or extreme environments for its high flexibility and exceptional light transmission capability[1, 2], including in vivo biomedical imaging[3] and internal inspection of expensive mechanical equipment such as jet engines and industrial pipelines[4].

4. We modified a sentence in the **Page 2 line 6** of the main text:

This limitation is exacerbated when attempting full-color imaging, where broadband chromatic aberrations inherent in the fiber bundle and imaging system distort projected contours and the point spread function (PSF).

5. We modified a sentence in the **Page 2 line 20** of the main text:

Such reciprocity exhibits on-demand Janus functionality[24-26], establishing a foundation for round-trip bifunctional full-color stereoscopic imaging system.

6. We modified a sentence in the **Page 3 line 1** of the main text:

Herein, we propose a Janus metafiber system for full-color stereoscopic visualization (FCSV) via a reciprocally bi-functional strategy, by fully combining forward and reverse optical paths through the fiber with the assistance of a metasurface-enabled quadratic phase, which simultaneously allows 2D spectral collection and high-precision 3D depth acquisition in an ultracompact footprint.

3. In the Abstract, the sentence beginning with "In this work, we propose …" is overly long. The term "multiplexing" is also used in a confusing way. The manuscript does not demonstrate "forward and reverse multiplexing" or "time division multiplexing." These buzzwords should be removed, and the methods should be described using clearer and more accurate terminology to improve accessibility.

Response: We thank the reviewer for raising this critical point. We have considered the accuracy of the term "multiplexing". In this work, we did not employ a strict rapid multiplexing configuration in time domain. Due to hardware limitations, we merely demonstrated the switching between forward and reverse directions using the identical optical path configuration. Consequently, the original description was not rigorous, and we have made corresponding revisions to several sentences in the main text.

Manuscript revision:

1. We modified a sentence in the **Page 1 line 6** of the main text:

Therefore, a large number of complex optical components are usually required to achieve dual-function imaging. In this work, we propose a single round-trip system with metasurface-integrated fiber bundle, namely the so-called Janus metafiber *system. With the forward and reverse switching of the same optical path, we simultaneously achieving passive 2D imaging in the visible regime and active 3D depth sensing.*

2. We modified a sentence in the **Page 1 line 10** of the main text:

The experiment shows high imaging fidelity and noise resilience benefiting from time-division acquisition between single-wavelength array scanning and polychromatic illumination.

3. We modified a sentence in the **Page 6 line 9** of the main text:

Given aforementioned endoscopic scheme, we use a quadratic phase metasurface to achieve reciprocal imaging, enabling to collimate incident light from all directions into a coaxial plane wave.

4. We modified a sentence in the **Page 18 line 14** of the main text:

By repeatedly switching the light path configuration, we simultaneously tested full-color 2D imaging and 3D depth imaging.

4. The SEM image in Figure 2f shows connected meta-atoms. Will this affect the PSF or overall imaging performance? Please clarify.

Response: Thank you for raising this important and insightful question. You have correctly identified a potential discrepancy between the fabricated structure and the design model. The minute connections between adjacent nanorods observed in the SEM image (**Fig. 2f**), are a common manifestation of nanofabrication edge roughness or proximity effects. We fully agree that, in principle, any geometric perturbation can alter the electromagnetic response of a unit cell. We would like to clarify this point from simulation and experimental aspects:

For the simulation setup, in our full-wave finite-difference time-domain (FDTD) simulations, we applied periodic boundary conditions (PBCs) along the x- and y-axes and perfectly matched layers (PMLs) along the z-axis. This setup carries a specific physical meaning: it models an infinite, perfectly periodic metasurface array, meaning the simulation automatically and precisely accounts for all electromagnetic coupling effects between adjacent nanorods under ideal periodic arrangement. The geometric parameters (length, width, height, rotation) of each nanorod we designed, and consequently the phase shift they provide, are the result of this strongly coupled system. Therefore, for the larger nanopillars clusters, the phase is likely accurate.

For the experimental performance, as shown in Fig. 3b, the measured PSF exhibits excellent agreement with the simulated PSF based on the ideal periodic model (in key metrics like main lobe width and side lobe levels). This directly proves that the overall phase profile has not undergone systematic deviation or severe distortion due to fabrication imperfections. For the final imaging performance, the imaging results presented in Fig. 3e and Fig. 5b clearly demonstrate that the metalens can successfully reconstruct target images with high quality and resolution approaching the diffraction limit. This is the most compelling evidence for the robustness of the system's performance.

Manuscript revision:

1. We added a sentence in the **Page 8 line 4** of the main text:

The minute connections between meta-atoms in Figure 2f represent common fabrication errors arising from small-scale exposure and etching processes, and do not significantly impact the optical performance of this device.

5. In Figure 3d, why does the PSF FWHM become smaller at a 10 degree

incidence angle?

Response: I really appreciate your comment on this problem. Under normal circumstances, the image plane formed by the secondary phase metasurface at the focal plane is formed along an ellipsoid[1] (when the object distance *d* is significantly greater than the focal length f), as shown in **Fig. R2a**. Therefore, considering the position of the imaging plane selected during PSF testing, even slight defocusing can cause deviation between the object plane of the test system and the clearly focused plane of the metasurface (**Fig. R2c**). During practical testing, to accommodate substantial focal plane displacement while maintaining the same field of view across all focal points, a 20X objective magnification was employed. At this point, the determination of the optimal position relative to the focal plane may exhibit slight deviations, potentially leading to a minor reduction in the full width at half maximum (FWHM) at a non-zero angle, as illustrated in **Fig. 3d** in the main text. However, this does not affect the subsequent two-dimensional imaging experiments (relying on the lens's inherent light field mapping capability) or the three-dimensional depth experiments (relying on the scanning intensity of the focused beam and the time-of-flight difference) that we demonstrate.

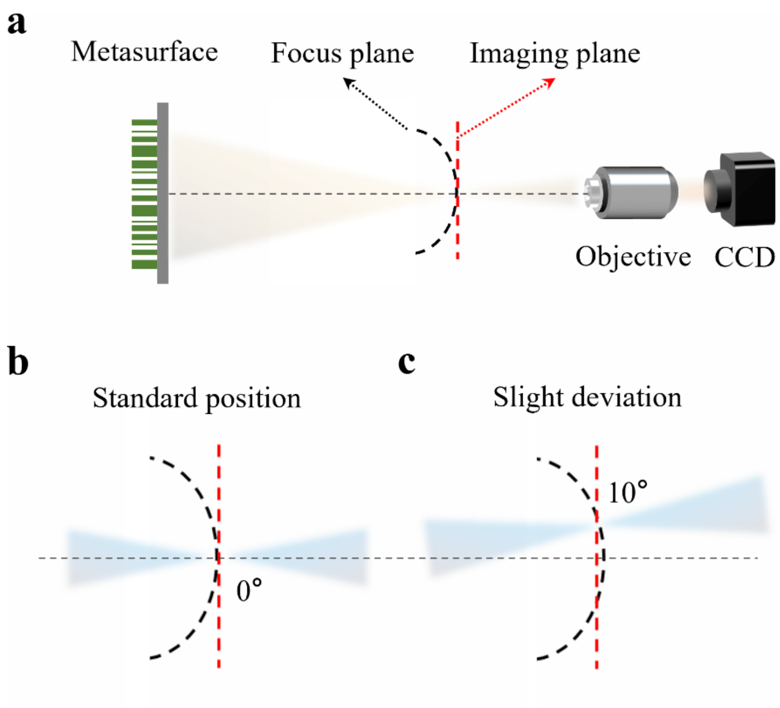

Figure R2. (a) The position relationship between the focus plane and the imaging plane. (b) The standard position between two planes. (c) The deviation position between two planes.

Manuscript revision:

1. We added a sentence in the **Page 9 line 21** of the main text:

As shown in the figure, the FWHM at the three wavelengths does not increase significantly within 0° to 30°. The small decrease of FWHM values at 10° is caused by the slight deviation between the focus plane and imaging plane.

6. In Figure 4f, why is the focusing efficiency higher at –10 degrees, and why is the response not symmetric around normal incidence?

Response: Thank you for your careful reading and correction. 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:

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

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. R3a**, 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. R3b**). The measured power P_f thus encompasses both the focused spot power and a portion of the transmitted background light power:

$$P_f = P_{focus} + P_{background}$$

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:

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

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.

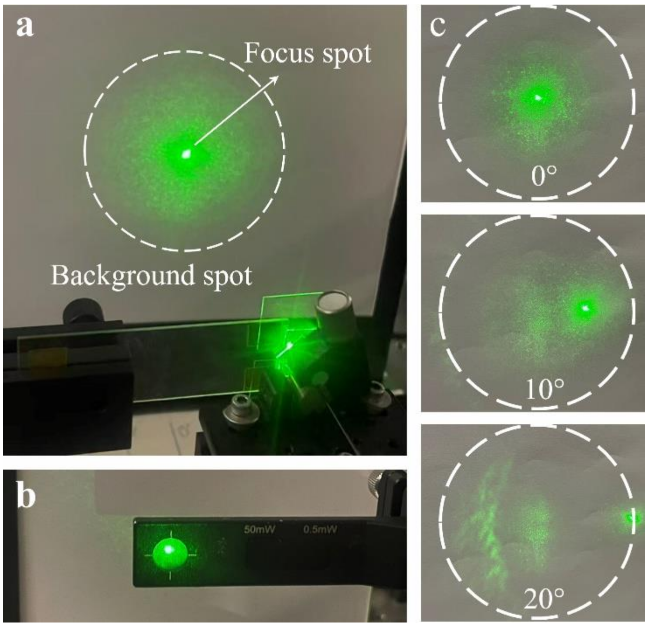

Figure R3. The spot component diagram of FCSV system.

Manuscript revision:

1. We added a sentence in the **Page 13 line 15** of the main text:

From the results, it can be observed that the focusing efficiency generally maintains constant (detailed analysis in Supporting Information Note 6 and Figure S8), and the transmittance and total device efficiency are decrease with increasing angle.

2. We added Supplementary Note 6, Fig. S8 in Supplementary information.

7. The purpose of the Gaussian fitting in Figures 4d and 4e is unclear. How does the green image on the left correspond to the Gaussian spot shown on the right in Figure 4d?

Response: We appreciate your comments on this problem. 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. R4a**:

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

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. R4b**). 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. R4c**. 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.

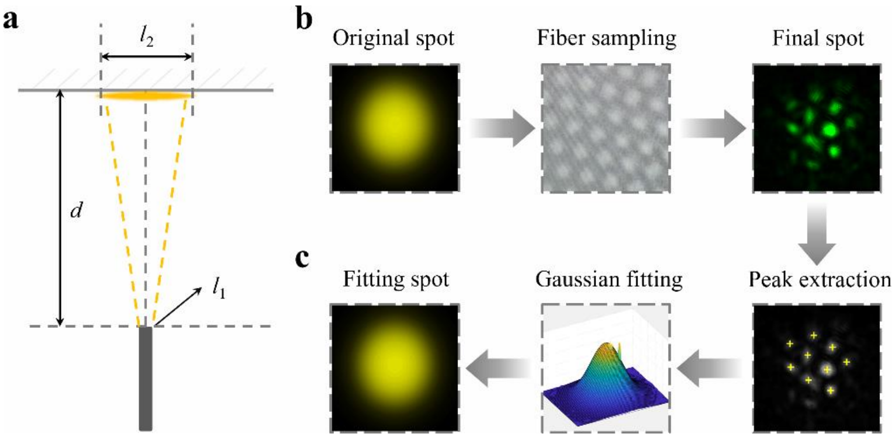

Figure R4. The 2D Gaussian fitting process of output spot.

Manuscript revision:

1. We added a sentence in the **Page 13 line 10** of the main text:

For estimating the original spot size to calculate the divergence angle, we used a two-dimensional Gaussian fitting function to fit the focused spots at different angles for calculating their FWHM, which is depicted in Supporting Information Note 5 and Figure S7.

2. We added Supplementary Note 5, Fig. S7 in Supplementary information.

8. The necessity and relevance of edge detection for fiber optic imaging applications should be discussed.

Response: We thank the reviewer for raising this important point. In fiber optic imaging systems, edge detection serves as the essential technological component throughout the entire workflow encompassing image pre-processing, feature extraction, and quantitative analysis.

- As fiber-optic imaging is affected by transmission losses, coupling efficiency, and environmental interference, issues such as low image contrast, blurred details, and excessive noise may arise. Edge detection techniques can be employed to enhance image recognition.
- In applications such as industrial inspection and biomedical imaging, precise target identification and quantitative measurement are required. Fiber-optic scanning imaging also necessitates layer-by-layer edge extraction of cross-sections to reconstruct three-dimensional surfaces.
- Concurrently, employing edge detection techniques enhances image feature recognition without hardware upgrades, thereby reducing computational load on backend processing.

Overall, employing edge detection within fiber optic imaging systems aligns with conventional imaging approaches, leveraging its inherent advantages of enhancing image recognition performance and reducing computational demands. Simultaneously, it enables the identification of critical components such as biological tissue contours, pipeline fissures, and mechanical structures within potential applications like fiber optic endoscopy. This renders imaging outcomes more multifunctional, reliable, and intelligent.

Manuscript revision:

1. We added the sentence in the **Page 18 line 8** of the main text:

In practical fiber-optic imaging applications, edge extraction techniques can be employed to enhance image contrast and reduce computational load, thereby *advancing the development of three-dimensional fiber-optic scanning imaging.*

9. The manuscript employs a simple metalens design without chromatic aberration correction, resulting in a smaller PSF at the design wavelength of 532 nm. The authors should clarify whether such a metalens is adequate for full colour imaging. The following paper should be cited in this context:

https://www.nature.com/articles/s41467-022-31902-3

Response: Thank you for the comments and suggestions. Chromatic aberration indeed poses a significant impediment to enhancing current 2D imaging performance, manifesting as slight degradation in PSF performance at non-target wavelengths, alongside multi-wavelength defocusing at the imaging plane and edge color separation. We have contemplated employing achromatic metalens to address this issue. However, current research on chromatic aberration mitigation faces substantial constraints, namely the mutual limitation between numerical aperture (NA) and diameter[2]. Even with smaller NA values, achromatic metalens typically achieve dimensions of merely tens of micrometers, which clearly fails to meet the imaging requirements of this work. Consequently, future research in this direction holds significant promise for achieving application-grade achromatic metasurface imaging. One potential avenue might be to explore the degrees of freedom available in nanocolumn design (such as z-axis height), enabling a broader range of phase-group delay dispersion combinations and thereby enhancing imaging performance.

Manuscript revision:

1. We added the sentence in the **Page 19 line 4** of the main text:

Chromatic aberration also frequently constitutes a significant constraint on full-color imaging performance. Future approaches may consider employing methods such as *extending the nanostructure height freedom*[31] *to achieve superior achromatic performance.*

Responses to Reviewer #2's comments

Reviewer #2: The manuscript presents an interesting metasurface-based fiber imaging system with promising experimental results. However, several key aspects of the system and its claimed performance require further clarification. The authors are advised to conduct a major revision before the manuscript is reconsidered.

Response: Thank you very much for the great support of our work. Your suggestions are very constructive, which will greatly help us to improve the manuscript. We carefully addressed all the issues you mentioned as follows and revised the paper.

1. Full-color 2D imaging and 3D ToF sensing rely on different illumination sources and detectors (LED/CCD versus laser/SLM/APD). Could the authors clarify whether both modalities can be realized through mode switching within a fixed system configuration, or whether changes in the optical layout or component positioning are required?

Response: We thank the reviewer for raising this critical point. The optical configuration diagram employed in the actual testing is presented in **Fig. R1**. 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.

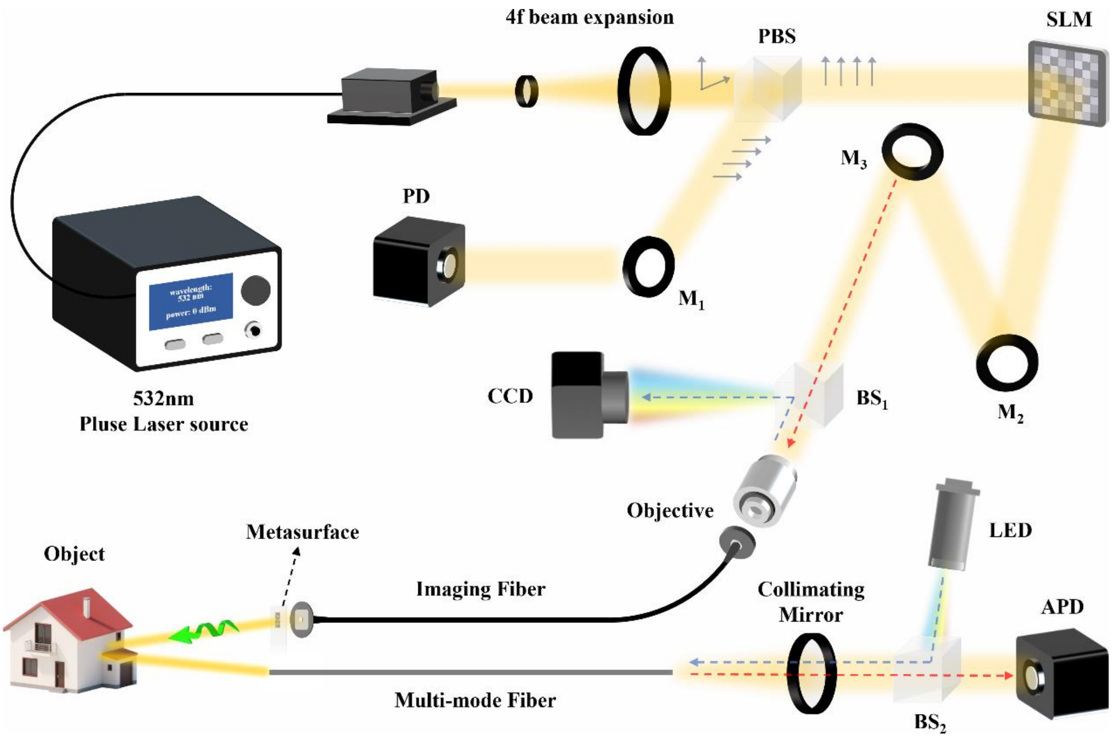

Figure R1. Optical configuration of FCSV.

Manuscript revision:

1. We added the sentence in the **Page 5 line 8** of the main text:

This is achieved by adding a beam splitter to route the forward and reverse incident signals and adopting a time-domain alternating acquisition strategy, which *sequentially acquiring full-color images and pulse signals and the interference between them can be avoided (shown in Supporting Information Note 1 and Figure S1).*

2. We added Supplementary Note 1, Fig. S1 in Supplementary information.

2. The manuscript introduces an additional multimode fiber for ToF illumination and signal collection. Could the authors elaborate on whether this MMF plays an essential role beyond light delivery and collection, and whether it offers clear advantages for system integration or miniaturization given the current architecture?

Response: We appreciate your comments on this problem. 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. R2**. 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.

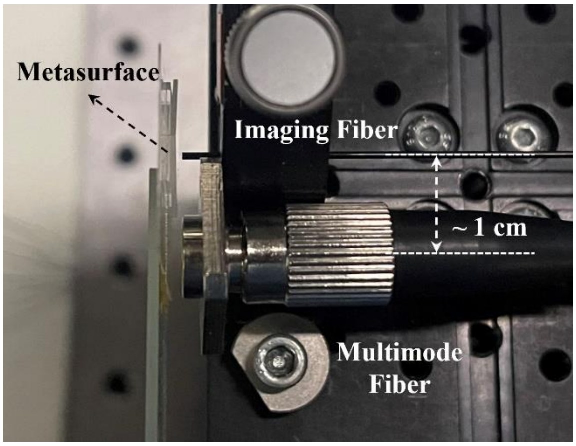

Figure R2. The diagram of the actual FCSV device.

Manuscript revision:

1. We added a sentence in the **Page 5 line 12** of the main text:

The actual picture of the output part of this system is shown in Supporting Information Note 2 and Figure S2).

2. We added Supplementary Note 2, Fig. S2 in Supplementary information.

3. As the metasurface appears to be optimized for a single working wavelength, could the authors clarify how broadband chromatic aberrations influence full-color imaging performance? In particular, how do the wavelength-dependent PSF variations observed in Fig. S2 affect image fidelity, and might dispersion-engineered metasurface designs help improve broadband focusing stability?

Response: We appreciate your constructive comment on this problem. Chromatic aberration stands as one of the core factors affecting imaging quality in full-colour imaging systems. Its fundamental cause lies in the differing refractive indices of optical media for visible light of varying wavelengths, preventing light of different colours from converging precisely onto the same focal plane. The impact of chromatic aberration manifests primarily in the following aspects:

- **Longitudinal chromatic aberration:** Light of different wavelengths passing through a lens focuses at different positions along the optical axis. For instance, blue light possesses a higher refractive index, resulting in a focal point closer to the lens; red light, with a lower refractive index, focuses further from the lens. This causes a situation where, on the imaging plane, one colour appears sharp while the other two colours are blurred. Consequently, the overall image sharpness diminishes, and fine details (such as edge textures and minute patterns) cannot be accurately reproduced. Simultaneously, colour shifts occur, saturation decreases, and the image's signal-to-noise ratio is reduced.
- **Lateral chromatic aberration:** Different wavelengths of light undergo varying degrees of magnification on the imaging plane, resulting in coloured fringing along image edges. For instance, during magnification of the watermelon image in **Figure 3e**, blue fringing appears along the edges, compromising line sharpness and directly diminishing resolution performance. This leads to inconsistent rendering between the scene's centre and periphery.

Specifically, as illustrated in **Fig. S4**, the wavelength-dependent PSF leads to the aforementioned spatially variable chromatic aberration, thereby precluding the simultaneous optimisation of resolution, contrast, saturation, and signal-to-noise ratio within the same plane. We have contemplated employing achromatic metalens to address this issue. However, current research on chromatic aberration mitigation faces substantial constraints, namely the mutual limitation between numerical aperture (NA) and diameter[1]. Even with smaller NA values, achromatic metalens typically achieve dimensions of merely tens of micrometers, which clearly fails to meet the imaging requirements of this work. Consequently, future research in this direction holds significant promise for achieving application-grade achromatic metasurface imaging. One potential avenue might be to explore the degrees of freedom available in nanocolumn design (such as z-axis height), enabling a broader range of phase-group delay dispersion combinations and thereby enhancing imaging performance.

Manuscript revision:

1. We added the sentence in the **Page 19 line 4** of the main text:

Chromatic aberration also frequently constitutes a significant constraint on full-color imaging performance. Future approaches may consider employing methods such as extending the nanostructure height freedom[31] to achieve superior achromatic performance.

4. The Introduction reports a field of view of 60°, whereas the main text focuses on performance within ±20° due to image degradation at larger angles. Could the authors clarify the distinction between the maximum geometrical FOV and the effective FOV with acceptable image quality?

Response: Thank you for your nice comments on our article. The theoretical limiting field of view (FOV) for this secondary phase metasurface is 60°, a characteristic basically 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. R3**. Due to **aperture clipping** at large angles and **background light attenuation** (as addressed in the response to Reviewer 1's Question 6), 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.

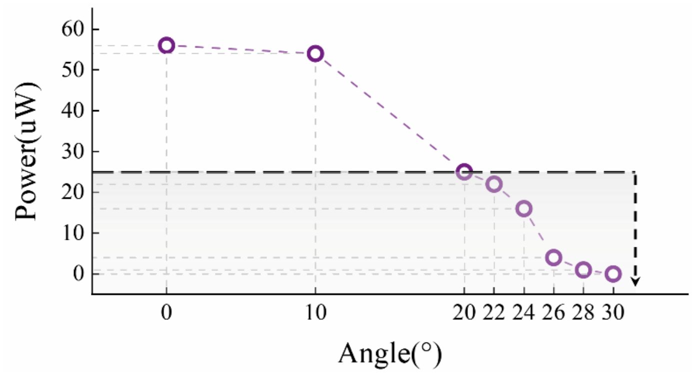

Figure R3. The power testing at different deflection angles.

Manuscript revision:

1. We added the sentence in the **Page 13 line 5** of the main text:

The image quality has deteriorated at the large angles owing to the limit of the edge of the aperture, and given the resolvable intensity of reflected light, so we will focus only on investigating the imaging performance between ± 20° for convenience. Detailed analysis is documented in Supporting Information Note 4 and Figure S6.

2. We added Supplementary Note 4, Fig. S6 in Supplementary information.

5. The reported improvement in depth resolution relies in part on interpolation techniques. Could the authors provide additional quantitative analysis to clarify how the claimed depth accuracy relates to the intrinsic system resolution?

Response: We thank the reviewer for raising this critical point. 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 Δ*t* of 1 ns). Assuming the scene depth is *d*, the minimum distinguishable depth difference is Δ*d*, 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. R4a**, there are:

$$\Delta t \cdot c = 2\Delta d$$

$$\Delta d = \frac{\Delta t \cdot c}{2} = \frac{1\times10^{-9} s \times 3\times10^{10} cm\cdot s^{-1}}{2} \approx 15\ cm$$

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. R4b**. At this time, there are:

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

Taking the interval *h* as 1 cm, the calculation yields Δ*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. R4c**. The maximum offset value

l_{max} of the emission point relative to the center point is taken as 2.5 cm. At this time, the relevant formula becomes:

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

After calculation, Δd = 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 ≈ 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.

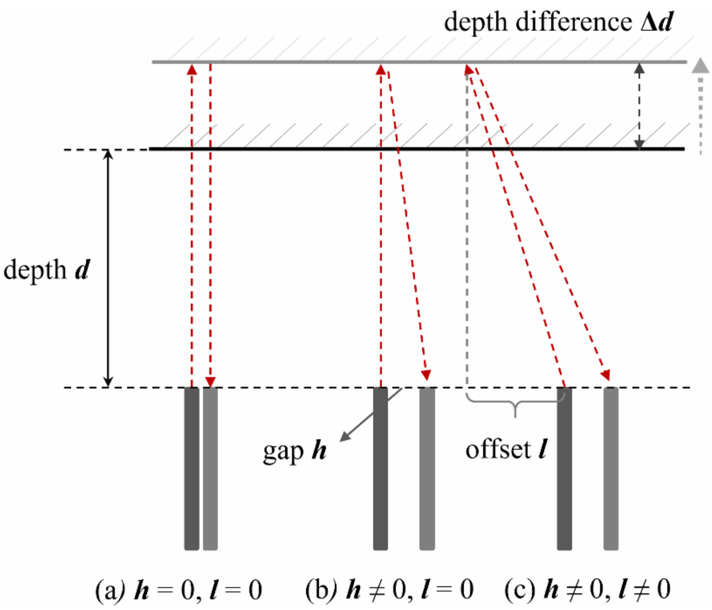

Figure R4. (a) Adjacent and normal emission situation. (b) Interval and normal emission situation. (c) Interval and oblique emission situation.

Manuscript revision:

1. We added the sentence in the **Page 14 line 6** of the main text:

As seen in the figure, clear resolution is achievable at five angles when selecting a 2 cm depth interval, representing over an eightfold improvement compared to the original accuracy, thereby validating the feasibility of this interpolation method. A detailed explanation of the intrinsic depth resolution and interpolation resolution of the system is demonstrated in the Supporting Information Note 8 and Figure S10.

2. We added Supplementary Note 8, Fig. S10 in Supplementary information.

6. A careful proofreading to address minor typographical and language issues would further improve the clarity of the manuscript.

Response: We really appreciate your careful reading of the manuscript. We have carefully double-checked the typos and grammatical inconsistencies in the main text as well as the Supplementary information. **All the changes have been marked in RED color** in the revised manuscript for your convenience.

Manuscript revision:

1. We revised the typos and grammatic mistakes in the main text as well as the Supplementary information. All the changes have been marked in RED color in the revised main text.

[1] E. Lassalle, T.W.W. Mass, D. Eschimese, A.V. Baranikov, E. Khaidarov, S. Li, R. Paniagua-Dominguez, A.I. Kuznetsov, Imaging Properties of Large Field-of-View Quadratic Metalenses and Their Applications to Fingerprint Detection, ACS Photonics 8 (2021) 1457-1468.

[2] Y. Wang, Q. Chen, W. Yang, Z. Ji, L. Jin, X. Ma, Q. Song, A. Boltasseva, J. Han, V.M. Shalaev, S. Xiao, High-efficiency broadband achromatic metalens for near-IR biological imaging window, Nat Commun 12 (2021) 5560.

Responses to Reviewer #1's comments

Reviewer #1: I appreciate the authors' careful attention to my previous comments. The revised manuscript now includes an optical image of the actual FCSV device, and I am satisfied that all technical issues have been addressed.

However, one critical concern remains: the integration of a metasurface with the fibre system. The current device does not demonstrate a true metafibre, which by definition involves direct integration of a metasurface onto the facet of an optical fibre, a concept well established in the literature. Since the manuscript claims "Full-color 3D Visualization with Janus Metafiber," this integration is essential to justify the central innovation and meet the expectations of Nature Communications.

I strongly encourage the authors to pursue a practical integration strategy that reflects the metafibre concept. This does not necessarily require direct fabrication on the fibre facet; even attaching a metasurface via a thin substrate can serve the purpose, provided the metasurface and fibre function as a unified device.

Without this integration, the technical advancement is significantly diminished, and the current implementation, relying on free-space and fibre-optic components, appears standard and below the threshold for publication in Nature Communications. With integration, however, I believe the manuscript is ready for acceptance.

Response: We sincerely appreciate the reviewers' insightful comments and constructive suggestions. 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 R1a**.

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 R1b**. 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 R1c** and **d**, where **Figure R1c** shows the integration of the lens group and imaging fiber through a capillary tube, and **Figure R1d** shows the image of integration sample of the secondary phase metasurface in this paper. In **Figure R1d**, 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.

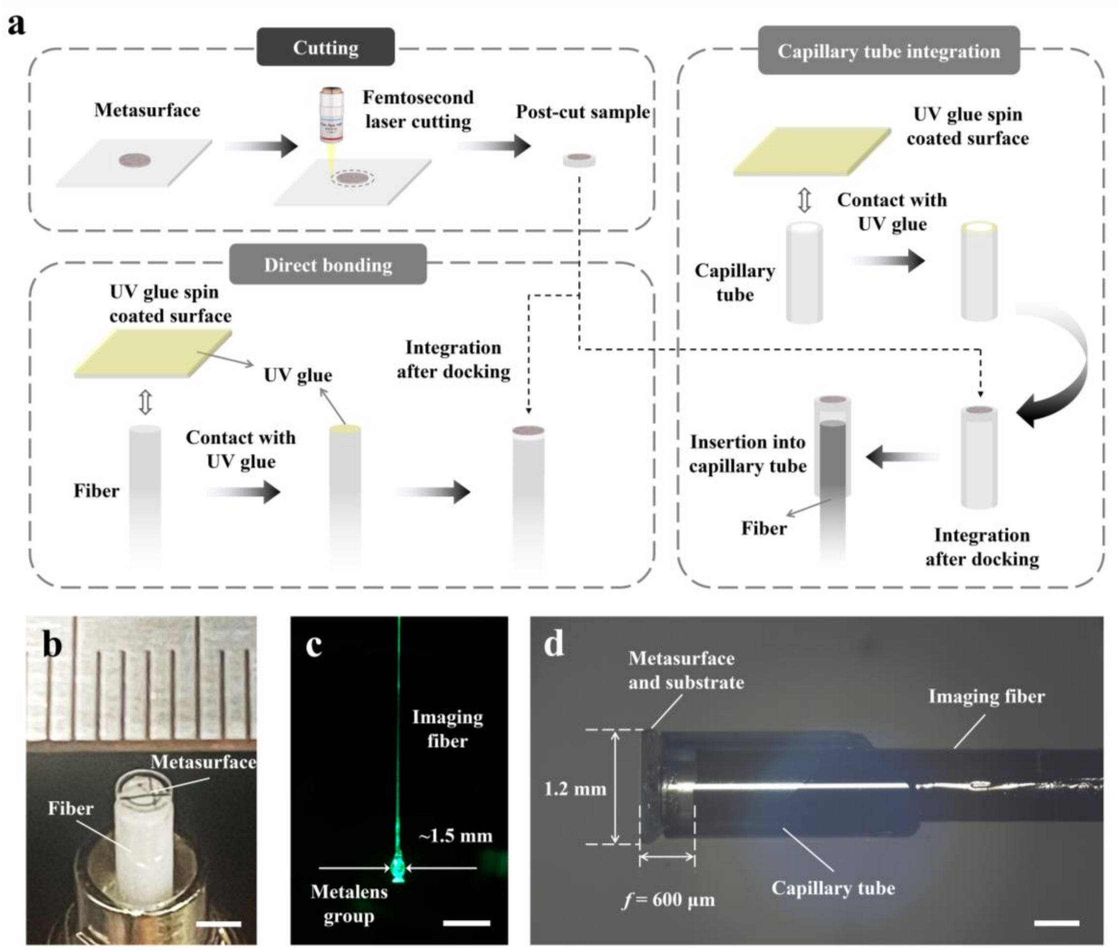

Figure R1. The integration strategy of metafiber system and related results. (a) The diagram of the integration process between metasurface and optical fiber. There are two main pathways: UV-curable adhesive assisted direct bonding and capillary tube–assisted alignment and packaging. **(b)** The image 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.

Manuscript revision:

1. We added a sentence in the **Page 18 line 20** of the main text:

Additionally, we demonstrated synthetic 3D full-color images optimized using multi-parameter optimization and edge detection results obtained through adjacent value determination method. For practical applications, metasurface samples must be robustly integrated with optical fibers to ensure mechanical stability, operational reliability, minimal overall footprint, and controlled the rigidity. The integration strategy of metafiber system has also been investigated and the related results are shown in Supporting Information Note 10 and Figure S12.

2. We added Supplementary Note 10, Fig. S12 in Supplementary information.

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