

Object Function Retrieval by Model-Based Optimization in Fourier Holographic Endoscopy

Thi Van Anh Nguyen,[†] Munkyu Kang,[†] Dinh Hoang Tran, Tran Thinh Le, Youngwoon Choi,* and Wonshik Choi*



Cite This: *ACS Photonics* 2024, 11, 385–394



Read Online

ACCESS |

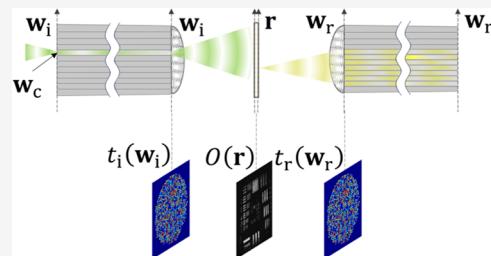
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Holographic endoscopes have garnered extensive attention not only for their compactness but also for their high-resolution and wide-field imaging capabilities. However, their practical use is challenging due to the reliance on Fresnel approximations in the underlying image reconstruction protocol, limiting their real-world applications. In this study, we present a model-based optimization algorithm that successfully retrieves an object function in a fiber bundle holographic endoscope. By developing a mathematical model that eliminates the need for approximations, our optimization extends the working range and expands the field of view of a fiber bundle holographic endoscope while preserving its advantages. These advancements significantly enhance the utility and feasibility of holographic endoscopes for various diagnostic procedures.

KEYWORDS: *holographic endoscopy, fiber bundle, model-based optimization, optical imaging, biophotonics*



INTRODUCTION

In various clinical diagnostics, endoscopes play a crucial role in examining the internal structures of organs that are not easily visualized through conventional imaging optics. Advancements in endoscopy over the past few decades have primarily focused on achieving high-resolution imaging, including at the subcellular level, to facilitate early detection of disease.^{1–4} Another important aspect of technical development for endoscopy is narrowing probe diameters, aiming to minimize patient discomfort without compromising resolution significantly. In addition, there is a growing need for high-resolution endoscopes that can operate without the use of external labeling agents to ensure optimal compatibility with the human body. Consequently, extensive efforts have been dedicated to developing thin, flexible, and label-free endoscopes while maintaining subcellular resolution. To pursue these goals, various innovations have emerged, such as the integration of endoscopes with confocal microscopy,^{4,5} or optical coherence tomography (OCT).^{6,7} Many of these techniques employ single-mode fibers, which offer the advantage of being resistant to bending and enabling a flexible operation. However, they often require scanning devices or additional optical components at the distal end of the fiber, resulting in an increase in the probe diameter. Therefore, the size of the endoscope probes is limited by the dimensions of the scanning elements attached to the fiber tips.^{8–10}

To eliminate the need for distal scanners and reduce the probe diameter, holographic endoscopes utilizing multimode fibers (MMFs) have been introduced. The inherent complexity of light transmission through MMFs, generating random

speckle patterns, is addressed by the concept of the transmission matrix (TM).^{11–15} By measuring the TM of the fiber, it becomes possible to compensate for complex mode mixing and restore high-resolution images.^{10,15–17} In addition to MMFs, fiber bundles are also commonly employed as imaging probes in holographic endoscopes.^{1,18} The TM approach effectively eliminates pixelation artifacts and core-dependent phase retardation of the fiber bundle, resulting in clean and continuous images with a resolution surpassing its core-to-core spacing.¹⁹ Moreover, by recording the complex field of light, holographic endoscopes have the capability to capture wide-field images in a lensless configuration without any optical or mechanical elements.^{15,20} This advancement enables us to reduce the endoscope size down to the dimension of the fiber probe. As a result, successful demonstrations of ultrathin endoscopes with dimensions as small as only a few hundred micrometers have been achieved. However, the TM method is highly sensitive to fiber conformation, making it suitable only for rigid-type endoscopes.

Significant efforts have been devoted to removing the requirement for measuring the TM to achieve a flexible endoscope without a distal scanner. One approach that has

Received: August 2, 2023

Revised: December 22, 2023

Accepted: December 22, 2023

Published: January 8, 2024



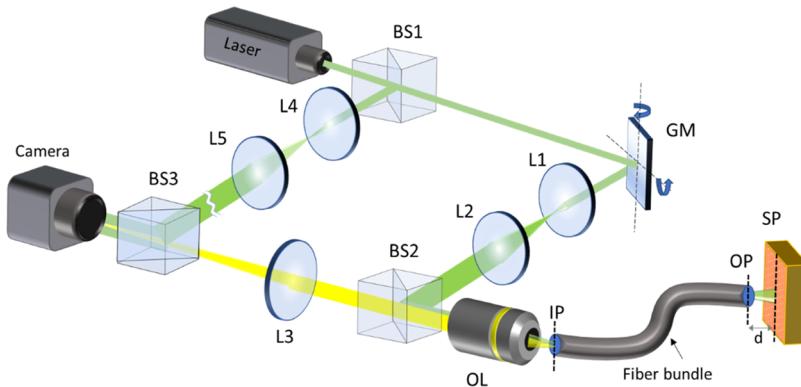


Figure 1. Experimental schematic of a Fourier holographic endoscope. GM: 2-axis galvanometer scanning mirror, BS1–3: beam splitters, L1–5: lenses, OL: objective lens. IP and OP: input and output planes of the fiber bundle, respectively. SP: sample plane. d : distance between the fiber tip and sample. The illumination beam and the reference beam are colored green, while the backscattered wave from the sample is colored yellow, although their wavelengths are all identical. The optical path length of the reference beam is matched to that of the sample beam within the coherence of the light source.

gained attention is the utilization of deep learning technology in holographic endoscopy.^{21–23} In this method, the TMs of the fibers are trained by using different conformations, allowing the inference of the TM for arbitrary conformations at a later stage. While promising, this approach has shown effectiveness primarily for simple artificial objects.^{24–26} Another approach involves the placement of a partial reflector at the distal end of a fiber bundle and using the reflection from this reflector as a reference wave to compensate for core-to-core phase retardations.²⁷

Recently, ultrathin Fourier holographic endoscopy operating in a reflection geometry has been demonstrated. In there, the light reflected from a target object through a fiber bundle was recorded in the form of a reflection matrix (RM).^{28,29} And the image reconstruction was performed by an algorithm termed the closed-loop accumulation of single scattering (CLASS),³⁰ where the phase retardations of the fiber bundle were dealt with as aberrations under the Fresnel approximation. The CLASS reconstruction enabled the extraction of an object image obscured by the core-dependent phase retardations of the fiber bundle. The RM measurement does not require prior calibration of the fiber bundle, allowing fully flexible endoscopic operations, regardless of the fiber conformation. Moreover, this method can image biological tissues without the need for labeling agents by effectively rejecting strong back reflections from the fiber's distal end. The CLASS reconstruction breaks down in scenarios where the Fresnel approximation ceases to be applicable. Its performance in image retrieval can be significantly compromised when the object is in close proximity to the fiber probe, resulting in a restricted working range or a narrow field of view (FOV).

In this study, we present a model-based optimization approach for lensless Fourier holographic endoscopy using a fiber bundle. We established a mathematical model that accurately describes the image acquisition process on a holographic endoscope. The model comprises three essential layers: input and output phase retardation maps accounting for core-dependent phase retardations of a fiber bundle and the object function representing the complex amplitude reflectance of the target object. Subsequently, we developed an optimization algorithm capable of identifying the three layers that maximize the similarity between the images measured from experiments and the images predicted by the model. By

relying on a precise description of light propagation instead of approximations, our method can accommodate extended working distances while maintaining a full FOV.

METHODS

Experimental Setup. Figure 1 depicts the experimental setup of our Fourier holographic endoscope system. A beam splitter (BS1) divides the output beam from a laser (either Finesse Pure, Laser Quantum, or custom-made Yb laser, wavelength: 1032 nm, frequency-doubled) into sample and reference beams. The sample beam is relayed to the back focal plane of an objective lens (OL, RMS10X, 0.25NA, Olympus) via a 4-f system consisting of two lenses (L1 and L2). A 2-axis galvanometer mirror (GM, GVS011/M, Thorlabs) placed at plane conjugate to the back focal plane of the OL steers the focal position of the sample beam at the input plane (IP) where a proximal end of a fiber bundle (FIGH-10-300S, Fujikura, diameter: 300 μm) is positioned. The fiber bundle consists of approximately 5500 fiber cores, each with a diameter of around 2 μm , and spans a length of 15 cm. Notably, the proposed method is not limited to this particular fiber length; rather, it is applicable to fiber bundles of any length. To illustrate this, we included imaging results with a 1 m long fiber bundle in the Supporting Information. The GM scans the sample beam in such a way as to scan the focused spot on each fiber core for delivering the beam to the distal end of the fiber bundle output plane (OP). After propagating through a free-space distance d , the sample beam reaches the sample plane (SP) and interacts with the sample. The backscattered light from the sample reached the distal end of the fiber bundle. It is then captured by the fiber bundle and delivered to the camera through the OL and a lens (L3). As the OL and L3 form a 4-f system, the proximal end of the fiber bundle is imaged on the camera. The reference beam is expanded by a 4-f telescope composed of two lenses (L4 and L5). The sample and reference beams meet at another beam splitter (BS3) at an angle and produce an off-axis interferogram with the camera. The path length of the reference beam is matched to that of the sample beam within the coherence length of the light source. By processing the interferograms, complex field maps of the sample are obtained. These complex field maps taken for illuminating at various fiber cores in the bundle serve as training data for the optimization algorithm,

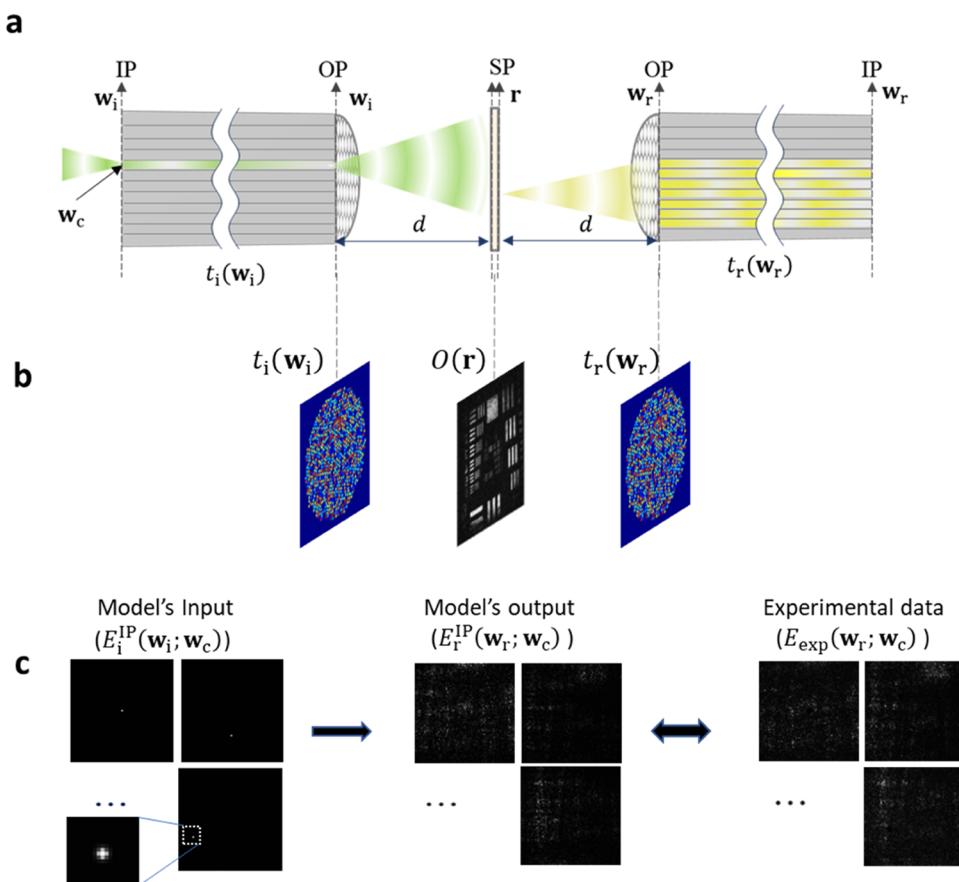


Figure 2. Light propagating model and generation of the reflected wave by the model. (a) Light propagating model. $t_i(\mathbf{w}_i)$: amplitude transmittance function of the fiber bundle in the illumination beam path. $O(\mathbf{r})$: complex amplitude reflectance of the target object. $t_r(\mathbf{w}_r)$: amplitude transmittance function of the fiber bundle in the reflection beam path. The illumination and reflection pathways have been unfolded to make their distinction clear. (b) Model layers made of $t_i(\mathbf{w}_i)$, $O(\mathbf{r})$, and $t_r(\mathbf{w}_r)$. (c) Data flow: point sources illuminating different fiber cores and reflected waves $E_r^{IP}(\mathbf{w}_r; \mathbf{w}_c)$ generated by the model. They are compared with experimental data $(E_{exp}(\mathbf{w}_r; \mathbf{w}_c))$ to find the model layers.

which yields the sample's object function and the phase retardation maps of the fiber bundle. The incident wave is set horizontally polarized, and only the vertical polarization component of the reflected wave is measured by placing a polarizer (not shown in Figure 1) in front of a camera. This cross-polarization detection eliminates reflection from the proximal side of the illumination core. Since there exists residual reflection from the distal side of the illumination core, the corresponding pixels are masked before processing the image. This cross-polarization detection scheme can be applied to most of the coherent imaging fiber bundles, as they are not made of polarization-maintaining fibers. In case the fibers maintain the polarization, then the temporal gating will play a major role in removing reflection from the proximal end. Since the coherence length of the laser used in our study (100 μm for a custom-made Yb laser or 6 mm for Finesse Pure, Laser Quantum) is much shorter than the fiber length, we can tune the reference beam path to selectively sample reflections from the target object.

Light Propagation Model. We set up a light propagation model that had the same configuration as the experiment. We then identified the object function by optimizing the similarity between the reflection images predicted by this model and the experimentally recorded images. Our model consists of three layers: the input amplitude transmittance function $t_i(\mathbf{w}_i)$ accounting for core-dependent phase retardations of the fiber bundle in the illumination beam path, the object function

$O(\mathbf{r})$, which represents the complex amplitude reflectance of the target object, and the amplitude transmittance function $t_r(\mathbf{w}_r)$ accounting for the fiber bundle in the reflection beam path (Figure 2a,b). t_i , t_r , and O are all complex-valued functions. Here, $\mathbf{w}_i = (u_i, v_i)$ is the lateral coordinate at the IP and OP for the illumination beam path, and $\mathbf{w}_r = (u_r, v_r)$ is the lateral coordinate for the reflection beam path. $\mathbf{r} = (x, y)$ represents the spatial coordinate of the SP. The distance between successive layers is d .

To generate the reflected wave by this light propagation model, we consider sending a focused wave to the fiber core located at a specific point $\mathbf{w}_c = (u_c^r, v_c^r)$ on the \mathbf{w}_i coordinate of the IP. While propagating through the fiber, the field experiences amplitude transmittance $t_i(\mathbf{w}_i)$ on its way to the OP. Thus, the field exiting the OP can be expressed as

$$E_i^{OP}(\mathbf{w}_i; \mathbf{w}_c) = t_i(\mathbf{w}_i)E_i^{IP}(\mathbf{w}_i; \mathbf{w}_c) \quad (1)$$

In our model, we assume $E_i^{IP}(\mathbf{w}_i; \mathbf{w}_c)$ to be a Gaussian wave focused by a lens with a numerical aperture of 0.25. The illumination wave exiting the fiber bundle propagates through a homogeneous medium to reach the SP and meets the target object as a spherical wave. The wave propagation between the layers is performed by the angular spectrum method. Specifically, we apply the amplitude transfer function of free space, $\tilde{H}(\mathbf{k}; z) = \exp[i\sqrt{(n\mathbf{k}_0)^2 - |\mathbf{k}|^2} d]$, to the spatial frequency spectrum of the incident wave within the passband

($|k| < nk_0$). Here, k is the transverse spatial frequency, $k_0 = 2\pi/\lambda$ is the free-space wavenumber, and n is the refractive index of the medium. Based on the angular spectrum method, the illumination wave in the SP can be written as

$$E_i^{\text{SP}}(\mathbf{r}; \mathbf{w}_c) = \mathcal{F}^{-1}\{\mathcal{F}\{E_i^{\text{OP}}(\mathbf{w}_i; \mathbf{w}_c)\}\tilde{H}(\mathbf{k}; d)\} \quad (2)$$

Here, $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ are the Fourier transform and inverse Fourier transform of the field, respectively. The illumination wave is reflected by the sample with the object function $O(\mathbf{r})$, which corresponds to the complex amplitude reflectance of the target sample. The reflected wave is then expressed as

$$E_r^{\text{SP}}(\mathbf{r}; \mathbf{w}_c) = O(\mathbf{r})E_i^{\text{SP}}(\mathbf{r}; \mathbf{w}_c) \quad (3)$$

In Figure 2a, this wave is depicted on the transmission side of the sample to differentiate it from the incident wave. The reflected wave travels back to the OP via another free-space propagation:

$$E_r^{\text{OP}}(\mathbf{w}_r; \mathbf{w}_c) = \mathcal{F}^{-1}\{\mathcal{F}\{E_r^{\text{SP}}(\mathbf{r}; \mathbf{w}_c)\}\tilde{H}(\mathbf{k}; d)\} \quad (4)$$

The reflected wave then returns to the IP while experiencing fiber-core-dependent amplitude transmittance $t_r(\mathbf{w}_r)$. After taking into account all of these processes, we obtain the reflected wave predicted by the model at the fiber proximal facet plane:

$$E_r^{\text{IP}}(\mathbf{w}_r; \mathbf{w}_c) = t_r(\mathbf{w}_r)E_r^{\text{OP}}(\mathbf{w}_r; \mathbf{w}_c) \quad (5)$$

This predicted wave $E_r^{\text{IP}}(\mathbf{w}_r; \mathbf{w}_c)$ by the model will be compared with the experimentally recorded field $E_{\text{exp}}(\mathbf{w}_r; \mathbf{w}_c)$ taken for the illumination at the same fiber core. More specifically, for each core illumination at \mathbf{w}_c , we have a model field E_r^{IP} at the camera determined by the three functions of t_i , t_r , and O . Then, our goal is to find these functions maximizing the similarity between the predicted model field and the experimentally recorded field at the camera for all \mathbf{w}_c .

Optimization. Our optimization algorithm is designed to find three layers in the model established in the previous section that minimize the loss function. The loss function for the optimization consists of the fidelity term L_f and regularization term R with the hyperparameter λ ranging from 0 to 1:

$$L = L_f + \lambda R \quad (6)$$

Here, L_f is used to evaluate how well the model fields fit the experimentally measured fields. In this study, we utilize two methods to calculate L_f : the mean squared error (MSE) and Pearson correlation. MSE is the most commonly used loss function for regression, and it calculates the mean squared differences between the true and predicted values for each data point. The MSE fidelity is given by

$$L_{\text{MSE}}(t_i, t_r, O) = \frac{1}{N_b^2} \sum_{\mathbf{w}_c}^{N_b} \sum_{\mathbf{w}_r} |E_{\text{exp}}(\mathbf{w}_r; \mathbf{w}_c) - E_r^{\text{IP}}(\mathbf{w}_r; \mathbf{w}_c)|^2 \quad (7)$$

$N_b = 30$ is the typical batch size for $N = 3000$ measured images. The MSE loss function is normalized such that its initial value is unity and its value is zero when the model fits perfectly the experimental data. MSE fidelity term serves as a baseline for comparison to the Pearson correlation method used in this study. In our experiment, the measured experimental fields contain noise that cannot be accounted for by our model. The examples include the stray reflections at

the fiber bundle and reflections from depths other than the depth of interest. To address the problems with a low signal-to-noise ratio, we employ the Pearson correlation coefficient L_p , which measures the similarity between the model data and experimental data:

$$L_p(t_i, t_r, O) = -\frac{1}{N_b} \times \left| \sum_{\mathbf{w}_c}^{N_b} \frac{\mathbb{E}[(E_r^{\text{IP}}(\mathbf{w}_r; \mathbf{w}_c) - \mu_r^{\text{IP}}(\mathbf{w}_c))(E_{\text{exp}}(\mathbf{w}_r; \mathbf{w}_c) - \mu_{\text{exp}}(\mathbf{w}_c))^*]}{\sigma_r^{\text{IP}}(\mathbf{w}_c)\sigma_{\text{exp}}(\mathbf{w}_c)} \right| \quad (8)$$

Here, $\mathbb{E}[\cdot]$ stands for the expectation value of the terms within the bracket with respect to \mathbf{w}_c , and $(\cdot)^*$ indicates the complex conjugation. $\mu_r^{\text{IP}}(\mathbf{w}_c)$ and $\sigma_r^{\text{IP}}(\mathbf{w}_c)$ are the mean and variance of $E_r^{\text{IP}}(\mathbf{w}_r; \mathbf{w}_c)$, respectively, and $\mu_{\text{exp}}(\mathbf{w}_c)$ and $\sigma_{\text{exp}}(\mathbf{w}_c)$ are similarly defined for $E_{\text{exp}}(\mathbf{w}_r; \mathbf{w}_c)$. When we evaluate the loss function, we take the summation with respect to the \mathbf{w}_c belonging to each batch. L_p is normalized such that its value is -1 when the model's output and experimental data are identical. As we shall show below, we find that the Pearson correlation method is effective in handling data with low signal-to-noise ratios, as it achieves faster convergence compared with using the MSE loss function.

To prevent overfitting during the optimization process, we introduced the regularization term R . L2-norm regularization is applied on the object function layer:

$$R = \sum_r |\mathcal{O}(\mathbf{r})|^2 \quad (9)$$

We employ the Adaptive Moment Estimation (ADAM) optimizer to optimize the complex-valued parameters t_i , t_r , and O for minimization of the loss function. Each of these complex parameters is decomposed into two real-valued components—one representing the real part and the other representing the imaginary part, for separate optimization. The algorithm is coded in Python, and TensorFlow version 1 serves as the framework for constructing the light propagating model and fine-tuning the model parameters.

In our experiment, we sent an incident wave with horizontal polarization and measured vertical polarization components to eliminate strong reflection from the proximal end of the fiber at the illumination core. In the context of optimization and image reconstruction, a lack of polarization may lead to the loss of information. However, the samples under consideration are mostly free from birefringence. Therefore, the scalar theory is a good approximation. Polarization-gated detection may lead to core-dependent amplitude attenuation. Precisely speaking, amplitude attenuation at each detection core may vary depending on the illumination core, and our algorithm finds their average value. t_i and t_r should be the same in theory. However, they are not identical due to experimental imperfection. There are uncontrollable phase fluctuations among complex field images taken for different illumination cores, which are added to the phase of t_i . For this reason, t_i and t_r should be optimized independently. The measured correlation between t_i and t_r was 0.004.

RESULTS

Comparison between Different Loss Functions. To explore the performance of the optimization, we used experimental data for the optimization with different loss functions. A United States Air Force (USAF) target (Edmund,

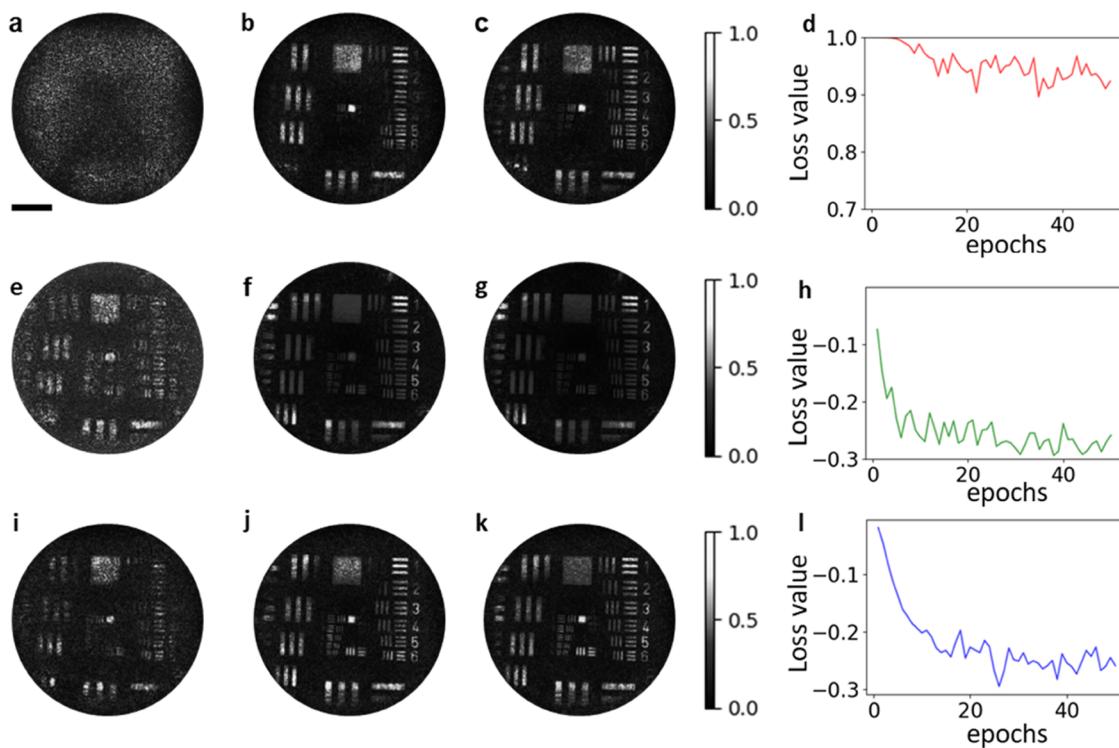


Figure 3. Comparison between different loss functions: (a–c), Object functions obtained by using the MSE loss function L_{MSE} after 1, 20, and 50 epochs with regularization. The standoff distance was $d = 200 \mu\text{m}$. Regularization hyperparameter $\lambda = 5 \times 10^{-3}$. (d) Loss function values with epochs. (e–h) Same as (a–d) but for Pearson correlation loss function L_p without regularization. (i–l) Same as (a–d) but for Pearson correlation loss function with regularization. Regularization hyperparameter $\lambda = 5 \times 10^{-4}$. Scale bars: 50 μm . Color bars: normalized amplitude.

2" x 2" Positive, 1951 USAF Hi-Resolution Target #58-198) was used as a test object. The target was placed on the SP at a fixed distance $d = 200 \mu\text{m}$ from the distal end of the fiber bundle at the OP. To detect the reflected wave from the target, we sent illumination light through a single fiber core and detected the returning wave through all of the fiber cores except for the illumination core. This is to reject the strong back reflection that can occur at the end surface of the fiber core. We acquired a total of 3000 reflection images with illumination through different cores, which takes 1.5 s when a fast camera with a 2000 Hz frame rate was used and 1 min for a 50 Hz camera. These interferograms were processed to obtain complex fields of reflection images. In the optimization process, the acquired reflection images were used to compare with the model output to retrieve the object function of the target.

Figure 3 presents the results obtained by using different loss functions with or without regularization. We conducted the optimization for 50 epochs with a batch size of 30. It takes 48 s for each epoch when the GPU of NVIDIA TITAN GTX is used. The Pearson loss function (Figure 3e–l) shows a convergence rate faster than that of the MSE loss function (Figure 3a–d). Furthermore, it provides better contrast and fidelity in finding the small structures (see the central parts of Figure 3c,g). In the case of the MSE loss function, its value decreases from 1 to 0.9 with optimization (Figure 3d), indicating that the light propagation model explains only 10% of the experimental data. On the other hand, for the Pearson correlation loss function in Figure 3l, the value decreases to -0.3 out of -1 , meaning that approximately 30% of the experimental data are accounted for by the model. Loss function values given in Figure 3d,h,l are computed using the

data from the last batch (30 images) in each epoch, which are used to optimize the model's parameters. Additionally, we computed the optimization errors across the entire data set of 3000 for the model's parameters at each epoch and found that they are in good agreement with the optimization errors of the last batch (see the Supporting Information). This supports that the random choice of the batch is a good representative of the entire data set.

With the optimization using Pearson loss, the existence of regularization leads to a faster and more stable convergence. In Figure 3i,j, the object functions exhibit less noise than in Figure 3e,f. Furthermore, the result of epoch 50 in Figure 3k with the regularization has a higher contrast, especially in the central part, compared to the result in Figure 3g with no regularization. Therefore, we used the Pearson correlation with the regularization as a loss function for the rest of analyses in this paper.

We compared t_r obtained for different loss functions (see the Supporting Information for details). For the subregion where target patterns were well identified, the correlation between Pearson loss with and without regularization was found to be 0.930, and the correlation between Pearson loss with regularization and MSE with regularization was 0.558. On the contrary, the correlation was much lower for the subregion where MSE loss does not work well.

Endoscopic Imaging of an Object at a Short Distance. We next investigated the performance of the optimization method for various standoff distances d and verified that the optimization method works at a short standoff distance where the Fresnel approximation is invalid. The effective FOV of our Fourier holographic endoscope based on the CLASS reconstruction is limited by the area in the SP that satisfies

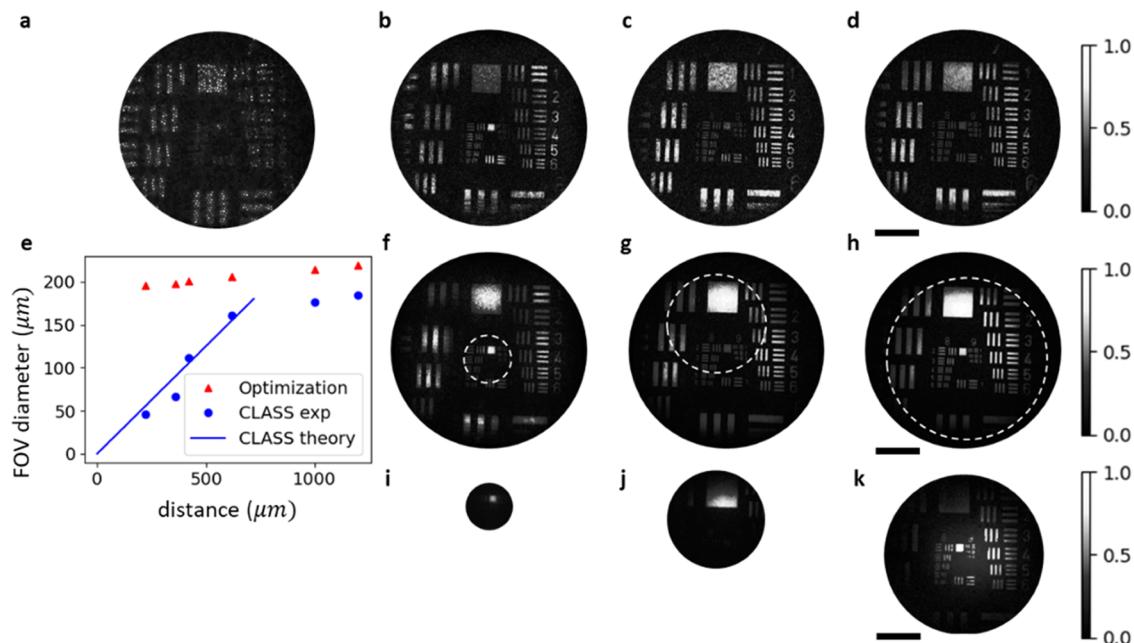


Figure 4. Comparison between the optimization method and the CLASS reconstruction for various standoff distances d . (a) Conventional endoscope image of the USAF target. (b–d) Object functions at standoff distances of 200, 400, and 600 μm in order, obtained by the optimization method using Pearson correlation with regularization as a loss function. (e) FOV diameters as a function of the standoff distance. (f–h) Synthesized image constructed by the object functions in (b–d). (i–k) CLASS images at the same distances as (b–d). The dashed circles in (f–h) indicate the FOV of the CLASS images at the corresponding distances. All scale bars: 50 μm . Color bars: normalized amplitude.

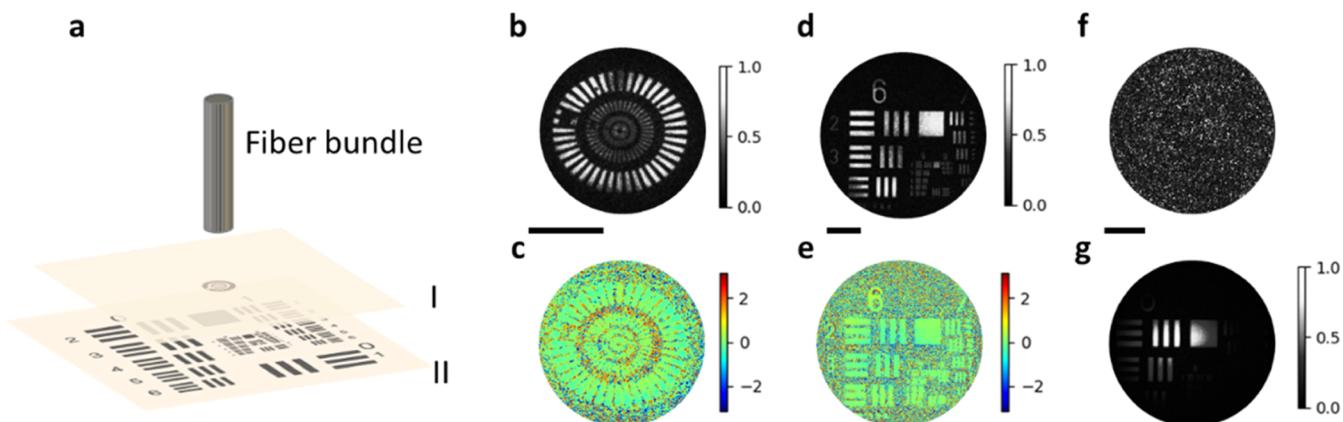


Figure 5. Three-dimensional (3D) image obtained by the optimization method. (a) Experimental configuration. Target I and target II are located at 210 and 850 μm from the tip of the fiber bundle, respectively. (b, c) Amplitude and phase maps of the object function of target I obtained by the optimization method. Pearson correlation with regularization was used as a loss function. (d, e) Same as (b, c) but for target II. (f, g) Images of targets I and II obtained by the CLASS reconstruction. All of the scale bars: 50 μm . Color scales in (b), (d), (f), and (g) were normalized by the maximum value in the corresponding maps. Color bars in (c) and (e): phase in radians.

the Fresnel approximation. Specifically, its view field diameter is determined by $L_{\text{eff}} = (\lambda/n)d\Delta D_{\text{eff}}^{-1}$, where ΔD_{eff} is the average core-to-core spacing of the fiber bundle.³¹ In applications where endoscopic imaging is conducted for targets with low reflectance, such as biological tissues, it often becomes necessary to position the fiber probe close to the sample to capture more signals. In such scenarios, the reduction of the view field diameter of the previous algorithms, which is proportional to the distance, becomes a major problem. The optimization algorithm proposed in this study uses the light propagation model based on the angular spectrum method. Therefore, no such limitation was imposed in the previous algorithm.

In Figure 4, we compare the reconstructed FOVs by the optimization method with those based on the previous CLASS reconstruction for various standoff distances. As a point of reference, we included a conventional endoscope image taken when the fiber bundle is in contact with the target object (Figure 4a). This image is discretized due to the core distribution of the fiber bundle. Figure 4b–4d presents the object functions of the target at standoff distances of 200, 400, and 600 μm , respectively, obtained by our optimization method. The FOV was decreased with the decrease of the standoff distance, but the view field diameter at a 200 μm distance was about 90% of that at a 600 μm distance. This slight reduction is mainly due to the reduction of the illumination area. In contrast, the FOV was substantially

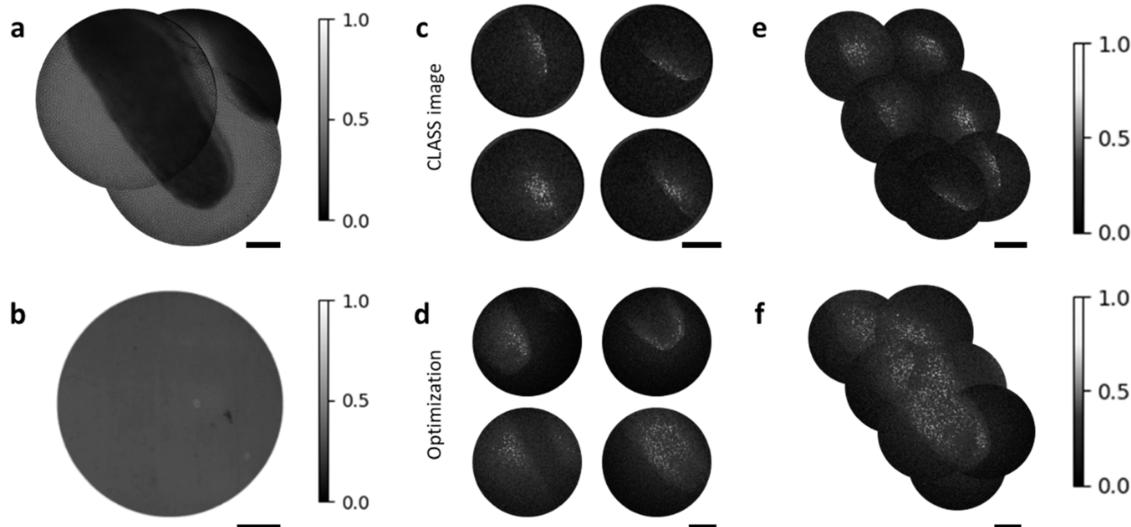


Figure 6. Imaging of a villus of a rat intestine. (a) Transmission LED image. (b) Image with a conventional endoscope. (c) Representative synthesized images by the CLASS reconstruction at various positions. (d) Representative synthesized images using the object functions retrieved by the optimization method. Pearson correlation with regularization was used as a loss function. (e) Combined image of those obtained by the CLASS reconstruction. (f) Combined image obtained by the optimization method. Scale bars: 50 μm . Color bars: normalized amplitude.

reduced with a decrease of the standoff distance for the CLASS reconstruction (Figure 4i–4k). The view field was reduced to 30% or less with the reduction of the standoff distance from 600 to 200 μm . The trend in the shorter standoff distance than 600 μm was in good agreement with the theoretical prediction based on Fresnel approximation (solid blue line in Figure 4e). For a larger standoff distance, the view field diameter converges to that of the optimization method, which is set by other factors such as the field of illumination.

It is important to note that the optimization process yields the sample's object function itself. However, most imaging methods, including CLASS reconstruction, produce object images that are affected by the transfer function of the imaging system. To compare the two methods on the same basis, we used the measured object functions in Figure 4b–4d to generate the object images by taking the transfer function of the fiber bundle into consideration. Specifically, we used the measured reflection matrix and synthesized the object images in the same way as the CLASS reconstruction after compensating the phase retardations by the fiber bundle with the maps identified by the optimization method. The resulting images are shown in Figure 4f–h. With the image synthesis, the overall contrast of the images is improved compared with the object functions shown in Figure 4b–4d. However, due to the uneven illumination, the edge areas of the images appear darker than the central parts. In contrast, when comparing the object functions obtained through our optimization method to those obtained through the CLASS reconstruction, our approach maintains the full FOV diameter consistently, regardless of the standoff distance. Moreover, the retrieved object functions exhibit uniform image contrast throughout the entire FOV.

3-Dimensional Imaging by a Single Set of Reflection Matrix Measurements. We prepared a bilayered sample that includes a Siemens star-like target (layer I) and a USAF target (layer II), separated by 640 μm (Figure 5a). We collected multiple reflection images from a fixed standoff distance of 210 μm from layer I while illuminating 3000 different cores. This

series of images, which forms a single reflection matrix measurement set, includes reflection signals from both layers.

The optimization method can reconstruct 3D images by scanning the standoff distance at each optimization. Figure 5b,c shows the amplitude and phase maps of the object function, respectively, obtained for layer I located at 210 μm . Similarly, Figure 5d,e shows those for layer II located at an 850 μm standoff distance. We confirmed the optimal training distance for each layer by observing the flat phase in the object functions, as shown in Figure 5c,e. If the distance used in the model is different from that of the experiment, there is a quadratic phase in the reconstructed object function associated with the discrepancy. Using the same data set, we attempted to obtain the object images at both layers by using the CLASS reconstruction (Figure 5f,g). At the long distance at layer II, the CLASS reconstruction produces the object image, but its FOV is significantly smaller than that of the optimization. The advantage of the optimization is particularly evident at the short distance of 210 μm , where the Fresnel approximation is not valid. While the CLASS reconstruction failed to resolve the Siemens star-like target (Figure 5f), the optimization yields a reconstruction quality enough to resolve the fine details (Figure 5b). Therefore, the optimization method presented in this study allows for 3D imaging capability with a single set of measurements, providing benefits of wider working ranges and larger FOVs compared to those of the CLASS reconstruction.

Once the fiber transmission is identified by an optimization at one depth, then it can be reused to retrieve the object function at the other depths. In fact, no further optimization is necessary, as we can apply the computational propagation to the measured complex field maps after correcting the fiber transmission to obtain object functions at an arbitrary depth from the initial depth. We demonstrated this in our earlier study based on the CLASS algorithm, and the same is the case for the present method. This can substantially reduce the computation time for 3D reconstruction. In this study, we scanned the standoff distance at each optimization. We intended to show that the scanning of the standoff distance is equivalent to scanning of the objective focus in 3D imaging.

The previous CLASS-based method does not have this capability, as the algorithm converges to a certain depth where the reflection signal is most prominent. However, scanning the standoff distance also has a limitation in the case when multiple layers are closely spaced. The optimization can converge to the layer, providing the strongest reflection signals near the standoff distance of the model. In such a case, we needed to employ the previous computational propagation method to reconstruct the volumetric image.

Imaging of Biological Samples. In this section, we evaluated the performance of our optimization algorithm for imaging of a biological specimen. As a sample, unstained mouse intestine tissue was mounted on an agarose gel plate and submerged in water to mitigate the refractive index contrast of the tissue. Before applying our method, we positioned one end of the fiber in direct contact with the intestine tissue. And, as a reference, we acquired both transmission and reflection images by utilizing a light-emitting diode (LED) light source. For the transmission image, the LED light was placed behind the tissue (Figure 6a). The shape of the villus is well visualized due to the significant contrast generated by tissue absorption. Subsequently, for the reflection image, we placed the LED light on the proximal side of the fiber while maintaining the contact of the fiber with the tissue. The light illumination and image acquisition were accomplished through the entire cores of the fiber bundle at the proximal end. The resulting reflection image, obtained using this conventional endoscope configuration, is presented in Figure 6b.

Figure 6c,d shows the results of individual imaging sites obtained using the CLASS reconstruction and optimization method, respectively. Due to the limited FOV of the CLASS reconstruction, the images in Figure 6c have a reconstruction diameter of about $100\ \mu\text{m}$. In contrast, the optimization offers a larger FOV of approximately $210\ \mu\text{m}$, resulting in significantly wider and smoother imaging coverage compared with the CLASS reconstruction. As shown in Figure 6e,f, the optimization produces more seamless stitching for different imaging sites, providing a comprehensive view of the entire imaging area.

■ DISCUSSION AND CONCLUSIONS

In this study, we have proposed a model-based optimization method for retrieving object functions in lensless Fourier holographic endoscopy using a fiber bundle. First, we developed a mathematical model that precisely describes the image acquisition process in our endoscopic configuration, considering three layers of light propagation. The first and third layers account for the core-dependent phase retardations that occur during the illumination and return of light through the fiber bundle, respectively. The second layer represents the reflection of light from the target object characterized by an object function represented as a complex amplitude reflectance map at a given standoff distance. By maximizing the similarity between experimentally acquired and model-predicted images, the optimization effectively identifies all three layers.

The proposed method can retrieve object images for any fiber configuration, as long as the shape of the fiber remains unchanged during image acquisition. This is because our approach does not require precalibration of the fiber bundle before imaging. We performed only the image acquisition sessions by measuring a set of reflection images for the scanning of the illumination cores, and the optimization

process separated the object function, input, and output phase retardations from the recorded images. In dynamic scenarios where the target object is continuously moving during a set of measurements, the quality of the reconstruction in our method is inevitably affected. In our previous demonstration,²⁸ we split the data into subsets and merge their results by exploiting the fact that the perturbation appeared as a linear phase ramp on the retardation maps. We will be able to employ the same strategy to counteract the image degradation even under dynamically relative motion between the fiber and sample.

Since the second layer models the complex amplitude reflectance of the target object, the optimization can successfully extract the object function itself. In most imaging methods, the object information is affected and modified by the transfer function of the imaging optics. Usually, the low-spatial-frequency band is more emphasized than the high-spatial-frequency region due to the transfer function with a triangular shape, resulting in a contrast falloff for fine features. Consequently, in the final image, the object is seen as if it is modified by a weighted low-pass filter, losing the contrast, particularly over the high-spatial-frequency region. On the contrary, by retrieving the object function as it is, the proposed optimization can produce accurate object information without being modified by the transfer function of the imaging system. Thus, the object information is acquired with a uniform contrast over the entire observable spatial frequency band.

Furthermore, the developed model utilizes the angular spectrum method, describing the light propagation between the layers accurately. Thus, our optimization produces a clear object function in any standoff distance, resulting in an extended operating range. The effect of building an accurate model becomes more dramatic with a short standoff distance when an object is in close proximity to the fiber bundle, showing that the achievable FOV is preserved almost the same. As a consequence, the lateral extent of the FOV is significantly larger than that of the CLASS reconstruction. As we reported earlier, the CLASS reconstruction based on the Fresnel approximation fails to work properly with an object too close to the imaging probe, causing a substantial narrowing of an FOV. However, since the optimization does not rely on such approximation, the achievable FOV is almost the same with a short standoff distance. Therefore, it has a significantly larger FOV than the CLASS reconstruction, particularly achieving about a 16 times increase at $200\ \mu\text{m}$.

It is also worth noting that there is a cost to pay for the advantages of our model-based optimization. The CLASS reconstruction exploits correlations among images to find and correct the core-dependent phase retardation by assuming the Fresnel approximation. Therefore, it solves the problem in a linear fashion, making image reconstruction considerably faster than the model-based optimization presented here. For instance, in the imaging of the resolution target shown in Figure 3, each epoch of optimization takes 48 s, and roughly 20 iterations are needed, while the CLASS reconstruction requires only 21 s per iteration and 10 iterations are needed, taking 4.6 times less. Therefore, the CLASS reconstruction is appropriate for a quick check of image results, while the optimization method is advantageous for postprocessing to obtain accurate images with a wider FOV.

In our light propagation model, we assumed that the core-to-core cross-talk is low enough to ignore. This means that when light is coupled to a certain core at the proximal end, light exclusively exits from the same core at the distal end and vice

versa. This assumption allows us to treat light transmission through a fiber bundle as an aberration problem, irrespective of the length of the fiber bundle, and describe its transmittance by $t_i(\mathbf{w}_i)$ and $t_o(\mathbf{w}_o)$, which are only the function of \mathbf{w}_i and \mathbf{w}_o , respectively. Given the notably low cross-talk ratio among cores in fiber bundles, this assumption has been largely valid. In case the cross-talk is substantial due to extreme bending and twisting, for instance, then the fiber transmission should be described by a transmission matrix $t_i(\mathbf{w}_i; \mathbf{w}_o)$ in the light propagation model. There also exists mode-to-mode cross-talk due to mode coupling. In our experiment, we selectively detect the lowest-order mode by means of time gating (see the Supporting Information) to minimize the effect of this cross-talk. In case the mode coupling is substantial in the detection scheme, then the fiber transmission should be described by the transmission matrix with respect to the modes.

In conclusion, our model-based optimization method, with its comprehensive mathematical model and accurate retrieval of object functions, extends the capabilities of a holographic endoscope using a fiber bundle while retaining its advantageous features such as fully flexible, ultrathin, and lensless configuration capable of 3D imaging of unstained biological tissues. It is apparent that the advantages of the optimization method we have demonstrated will make holographic endoscopes more effective in clinical practice. The wider working range mitigates the challenge of preparing a proper imaging condition, thereby facilitating the use of holographic endoscopes for the diagnostic process of abnormalities or lesions. Moreover, the larger FOV offered by our optimization reduces the need for extensive navigation of the endoscopic probe. With these significant advantages, our approach can contribute to the enhanced utility and feasibility of holographic endoscopes for various diagnostic procedures in real-world applications.

ASSOCIATED CONTENT

Data Availability Statement

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

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsphtronics.3c01100>.

Results with 1 m long fiber bundle; effect of batch size on resulting object function and loss value of optimization method; Pearson correlation value for entire data set at each epoch; amplitude and phase of reconstructed t_i , t_o , and O ; selection of lowest-order mode; determination of the training distance for optimization; and spatial resolution of the endoscopic imaging (PDF)

AUTHOR INFORMATION

Corresponding Authors

Youngwoon Choi — Department of Bioengineering and Interdisciplinary Program in Precision Public Health, Korea University, Seoul 02841, Republic of Korea; orcid.org/0000-0001-8492-8225; Email: youngwoon@korea.ac.kr

Wonshik Choi — Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 02841, Republic of Korea; Department of Physics, Korea University,

Seoul 02841, Republic of Korea; orcid.org/0000-0002-9445-7081; Email: wonshik@korea.ac.kr

Authors

Thi Van Anh Nguyen — Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 02841, Republic of Korea; Department of Physics, Korea University, Seoul 02841, Republic of Korea

Munkyu Kang — Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 02841, Republic of Korea; Department of Physics, Korea University, Seoul 02841, Republic of Korea; orcid.org/0000-0002-3645-5467

Dinh Hoang Tran — Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 02841, Republic of Korea; Department of Physics, Korea University, Seoul 02841, Republic of Korea

Tran Thinh Le — Center for Molecular Spectroscopy and Dynamics, Institute for Basic Science (IBS), Seoul 02841, Republic of Korea; Department of Physics, Korea University, Seoul 02841, Republic of Korea

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acsphtronics.3c01100>

Author Contributions

[†]T.V.A.N. and M.K. contributed equally to this work.

Funding

Institute for Basic Science (IBS-R023-D1); National Research Foundation of Korea (NRF-2021R1A2C2012069).

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Flusberg, B. A.; Cocker, E. D.; Piyawattanametha, W.; Jung, J. C.; Cheung, E. L. M.; Schnitzer, M. J. Fiber-Optic Fluorescence Imaging. *Nat. Methods* **2005**, *2* (12), 941–950.
- (2) De Groen, P. C. History of the Endoscope [Scanning Our Past. *Proc. IEEE* **2017**, *105* (10), 1987–1995.
- (3) Zhong, W.; Celli, J. P.; Rizvi, I.; Mai, Z.; Spring, B. Q.; Yun, S. H.; Hasan, T. In Vivo High-Resolution Fluorescence Microendoscopy for Ovarian Cancer Detection and Treatment Monitoring. *Br. J. Cancer* **2009**, *101* (12), 2015–2022.
- (4) Udovich, J.; Rouse, A.; Tanbakuchi, A.; Brewer, M.; Sampliner, R.; Joshua Udovich, al A.; Rouse, A. R.; Brewer, M. A. Confocal Microendoscope for Use in a Clinical Setting. In *Endoscopic Microscopy II*; SPIE, 2007; Vol. 6432, pp 81–89.
- (5) Jabbour, J. M.; Saldua, M. A.; Bixler, J. N.; Maitland, K. C. Confocal Endomicroscopy: Instrumentation and Medical Applications. *Ann. Biomed. Eng.* **2012**, *40* (2), 378–397.
- (6) Adler, D. C.; Chen, Y.; Huber, R.; Schmitt, J.; Connolly, J.; Fujimoto, J. G. Three-Dimensional Endomicroscopy Using Optical Coherence Tomography. *Nat. Photonics* **2007**, *1* (12), 709–716.
- (7) Li, X.; Gora, M. J.; Suter, M. J.; Tearney, G. J. Endoscopic Optical Coherence Tomography: Technologies and Clinical Applications [Invited]. *Biomed. Opt. Express* **2017**, *8* (5), 2405–2444.
- (8) Seibel, E. J.; Brown, C. M.; Dominitz, J. A.; Kimmy, M. B. Scanning Single Fiber Endoscopy: A New Platform Technology for Integrated Laser Imaging, Diagnosis, and Future Therapies. *Gastrointest. Endosc. Clin. N. Am.* **2008**, *18* (3), 467–478.
- (9) Liang, W.; Hall, G.; Messerschmidt, B.; Li, M. J.; Li, X. Nonlinear Optical Endomicroscopy for Label-Free Functional Histology in Vivo. *Light: Sci. Appl.* **2017**, *6* (11), No. e17082.
- (10) Dickensheets, D. L.; Kino, G. S. Silicon-Micromachined Scanning Confocal Optical Microscope. *J. Microelectromech. Syst.* **1998**, *7* (1), 38–47.

- (11) Moser, C.; Farahi, S.; Papadopoulos, I.; Loterie, D.; Goy, A.; Psaltis, D. Digital Confocal Microscopy through a Multimode Fiber. *Opt. Express* **2015**, *23* (18), 23845–23858.
- (12) Mahalati, R. N.; Kahn, J. M.; Gu, R. Y. Design of Flexible Multi-Mode Fiber Endoscope. *Opt. Express* **2015**, *23* (21), 26905–26918.
- (13) Niv, E.; Piestun, R.; Caravaca-Aguirre, A. M.; Conkey, D. B. Real-Time Resilient Focusing through a Bending Multimode Fiber. *Opt. Express* **2013**, *21* (10), 12881–12887.
- (14) Popoff, S. M.; Lerosey, G.; Carminati, R.; Fink, M.; Boccara, A. C.; Gigan, S. Measuring the Transmission Matrix in Optics: An Approach to the Study and Control of Light Propagation in Disordered Media. *Phys. Rev. Lett.* **2010**, *104* (10), No. 100601.
- (15) Choi, Y.; Yoon, C.; Kim, M.; Yang, T. D.; Fang-Yen, C.; Dasari, R. R.; Lee, K. J.; Choi, W. Scanner-Free and Wide-Field Endoscopic Imaging by Using a Single Multimode Optical Fiber. *Phys. Rev. Lett.* **2012**, *109* (20), No. 203901.
- (16) Kino, G. S.; Dickensheets, D. L. Micromachined Scanning Confocal Optical Microscope. *Opt. Lett.* **1996**, *21* (10), 764–766.
- (17) Orth, A.; Ploschner, M.; Wilson, E. R.; Maksymov, I. S.; Gibson, B. C. Optical Fiber Bundles: Ultra-Slim Light Field Imaging Probes. *Sci. Adv.* **2019**, *5* (4), No. eaav155.
- (18) Oh, G.; Chung, E.; Yun, S. H. Optical Fibers for High-Resolution In Vivo Microendoscopic Fluorescence Imaging. *Opt. Fiber Technol.* **2013**, *19* (6), 760–771.
- (19) Kim, D.; Moon, J.; Choi, W.; Kim, M.; Yang, T. D.; Chung, E.; Kim, J. Toward a Miniature Endomicroscope: Pixelation-Free and Diffraction-Limited Imaging through a Fiber Bundle. *Opt. Lett.* **2014**, *39* (7), 1921–1924.
- (20) Orth, A.; Ploschner, M.; Wilson, E. R.; Maksymov, I. S.; Gibson, B. C. Optical Fiber Bundles: Ultra-Slim Light Field Imaging Probes. *Sci. Adv.* **2019**, *5* (4), No. eaav1555, DOI: [10.1126/sciadv.aav1555](https://doi.org/10.1126/sciadv.aav1555).
- (21) Rahmani, B.; Loterie, D.; Konstantinou, G.; Psaltis, D.; Moser, C. Multimode Optical Fiber Transmission with a Deep Learning Network. *Light: Sci. Appl.* **2018**, *7* (1), No. 69.
- (22) Moser, C.; Kakkava, E.; Borhani, N.; Psaltis, D. Learning to See through Multimode Fibers. *Optica* **2018**, *5* (8), 960–966.
- (23) Resisi, S.; Popoff, S. M.; Bromberg, Y. Image Transmission Through a Dynamically Perturbed Multimode Fiber by Deep Learning. *Laser Photonics Rev.* **2021**, *15* (10), No. 2000553.
- (24) Kuschmierz, R.; Scharf, E.; Ortega-González, D. F.; Glosemeyer, T.; Czarske, J. W.; Kuschmierz, R.; Scharf, E.; Ortega-González, D. F.; Glosemeyer, T.; Czarske, J. W. Ultra-Thin 3D Lensless Fiber Endoscopy Using Diffractive Optical Elements and Deep Neural Networks. *Light: Adv. Manuf.* **2021**, *2* (4), 415–424.
- (25) Sun, J.; Wu, J.; Koukourakis, N.; Cao, L.; Kuschmierz, R.; Czarske, J. Real-Time Complex Light Field Generation through a Multi-Core Fiber with Deep Learning. *Sci. Rep.* **2022**, *12* (1), No. 7732.
- (26) Wu, J.; Wang, T.; Uckermann, O.; Galli, R.; Schackert, G.; Cao, L.; Czarske, J.; Kuschmierz, R. Learned End-to-End High-Resolution Lensless Fiber Imaging towards Real-Time Cancer Diagnosis. *Sci. Rep.* **2022**, *12* (1), No. 18846.
- (27) Badt, N.; Katz, O. Real-Time Holographic Lensless Micro-Endoscopy through Flexible Fibers via Fiber Bundle Distal Holography. *Nat. Commun.* **2022**, *13* (1), No. 6055.
- (28) Choi, W.; Kang, M.; Hong, J. H.; Katz, O.; Lee, B.; Kim, G. H.; Choi, Y.; Choi, W. Flexible-Type Ultrathin Holographic Endoscope for Microscopic Imaging of Unstained Biological Tissues. *Nat. Commun.* **2022**, *13* (1), No. 4469.
- (29) Kang, M.; Kang, M.; Choi, W.; Choi, W.; Choi, W.; Choi, W.; Choi, W.; Choi, Y.; Choi, Y.; Choi, Y. Fourier Holographic Endoscopy for Imaging Continuously Moving Objects. *Opt. Express* **2023**, *31* (7), 11705–11716.
- (30) Kang, S.; Kang, P.; Jeong, S.; Kwon, Y.; Yang, T. D.; Hong, J. H.; Kim, M.; Song, K. D.; Park, J. H.; Lee, J. H.; Kim, M. J.; Kim, K. H.; Choi, W. High-Resolution Adaptive Optical Imaging within Thick Scattering Media Using Closed-Loop Accumulation of Single Scattering. *Nat. Commun.* **2017**, *8* (1), No. 2157.
- (31) Choi, W.; Kang, M.; Hong, J. H.; Katz, O.; Lee, B.; Kim, G. H.; Choi, Y.; Choi, W. Flexible-Type Ultrathin Holographic Endoscope for Microscopic Imaging of Unstained Biological Tissues. *Nat. Commun.* **2022**, *13* (1), No. 4469, DOI: [10.1038/s41467-022-32114-5](https://doi.org/10.1038/s41467-022-32114-5).