



# Time-averaged image projection through a multimode fiber

DIRK BOONZAJER FLAES,<sup>1,\*</sup>  HANA ŠTOLZOVÁ,<sup>2</sup> AND TOMÁŠ ČIŽMÁR<sup>1,2,3</sup> 

<sup>1</sup>Leibniz Institute of Photonic Technology, Albert-Einstein-Straße 9, 07745 Jena, Germany

<sup>2</sup>Institute of Scientific Instruments of CAS, Královopolská 147, 612 64 Brno, Czech Republic

<sup>3</sup>Institute of Applied Optics, Friedrich Schiller University Jena, Fröbelstieg 1, 07743 Jena, Germany

\*Dirk.Boonzajer@leibniz-ipht.de

**Abstract:** Many disciplines, ranging from lithography to opto-genetics, require high-fidelity image projection. However, not all optical systems can display all types of images with equal ease. Therefore, the image projection quality is dependent on the type of image. In some circumstances, this can lead to a catastrophic loss of intensity or image quality. For complex optical systems, it may not be known in advance which types of images pose a problem. Here we show a new method called Time-Averaged image Projection (TAP), allowing us to mitigate these limitations by taking the entire image projection system into account despite its complexity and building the desired intensity distribution up from multiple illumination patterns. Using a complex optical setup, consisting of a wavefront shaper and a multimode optical fiber illuminated by coherent light, we succeeded to suppress any speckle-related background. Further, we can display independent images at multiple distances simultaneously, and alter the effective sharpness depth through the algorithm. Our results demonstrate that TAP can significantly enhance the image projection quality in multiple ways. We anticipate that our results will greatly complement any application in which the response to light irradiation is relatively slow (one microsecond with current technology) and where high-fidelity spatial distribution of optical power is required.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

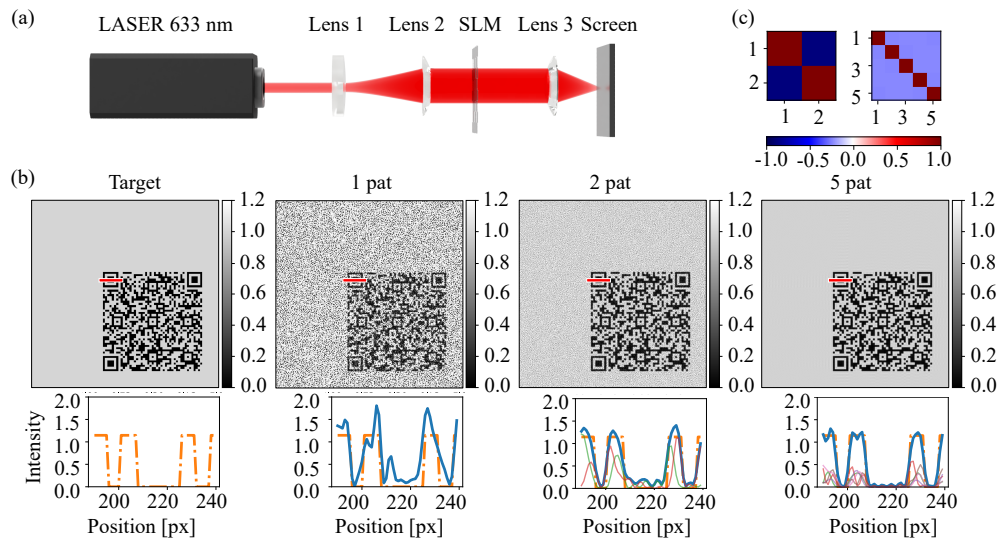
Applications of modern photonics often require three-dimensional structured light delivery in a hard-to-reach or sensitive environment, such as inside a live brain. In optical image projection systems, not all intensity distributions can be displayed with equal ease, and which intensity distributions are well-matched to the imaging system may not even be known exactly *a priori*, for instance when projecting images through a complex medium such as through a layer of paint or a multimode optical fiber [1–3]. We developed a general method, which we call Time-Averaged image Projection (TAP), for image projection through a complex optical system with optimal control over the intensity distribution. This algorithm relies on an image projector, such as a phase-only spatial light modulator (SLM) or a digital micro-mirror device (DMD), with a high frame rate. Instead of attempting to find a single pattern on the wavefront shaper that results in the required intensity, our algorithm decomposes the target intensity into a set of intensities that on average yield an intensity distribution that matches the target intensity distribution.

Our method allows for unprecedented image projection quality, avoiding known limitations in our optical setup, such as the projection of a homogeneous background. Moreover, multiple images can be displayed simultaneously at different depths. This can be tuned to extend the depth of focus, or to make it artificially narrow. Surprisingly, extending the depth of focus can be done with no discernible quality drop in the image projection. This would be very difficult to accomplish in cases where only a single pattern can be used.

Image projection through a multimode fiber has been demonstrated before in two dimensions, and a number of algorithms have been developed for image projection on the distal end of

a complex medium, either based on fully measuring the characteristics of the optical system by measuring the Transmission Matrix (TM) directly [2,4–11], or based on a neural network architecture [12,13]. In all these applications, a single pattern was used to render the entire scene and care must be taken to avoid strong speckle-like artifacts. Alternatively, images were generated by rapidly scanning diffraction-limited points over the distal end of a fiber [14] in a similar way to modern pico-projectors. This method is computationally much less demanding and it does not suffer from speckle-like artifacts, but displaying uniform backgrounds requires many scanning points. Time-averaged intensity patterns have been successfully employed for microscopy [15] and in optogenetics [16], but in those cases a rotating diffuser was employed. When direct optical access to the area in which the image is projected is not available, such as at the distal end of a multimode fiber, this would not be possible.

A proof of concept is shown in Fig. 1. We consider a theoretical setup in which the Fourier plane of a phase-only SLM is imaged onto a screen (Fig. 1(a)). This SLM can alter the spatial phase retardation of the beam, but not its amplitude. In a ray optics analogy, it can redirect light but not attenuate the beam. This setup is well-matched to display a focal point somewhere on the screen, as it only requires a linear phase ramp on the SLM. Generating a scene with uniform intensity in the imaging plane is not very well matched to this optical system, as it would require the light to originate only from a single spot on the SLM. Therefore, when the SLM is uniformly illuminated, almost all incident light is not usable, and unless it is somehow removed, it will only deteriorate the image projection quality. This is precisely the problem that our approach intends to solve.



**Fig. 1.** Proof of concept of time-averaged projection (TAP). We conceptually visualize the performance of a setup (a) in which the Fourier plane of an SLM is imaged onto a screen. (b) Target image and (from left to right) resulting intensity profile for 1, 2, and 5 patterns per frame. Bottom row: cross-section of the optimized results. The cross-sectional area is indicated by a red bar in the corresponding images. Orange dashed line: target image. Blue line: resulting intensity distribution. Thin lines correspond to the resulting intensities of individual patterns. (c) Normalized cross-correlation between the target image and the resulting intensity of each individual pattern. The cross-correlation of the off-diagonal elements is negative, indicating that the patterns are indeed complementary.

In our simulation, we intend to find a set of patterns to be displayed on the SLM that result in a uniform bright area with a Quick Response (QR) code in the lower right of the screen. This is a challenging target image for such an optical system, as it both incorporates uniform illumination and high-frequency features. When we use only a single pattern on the SLM to generate the frame, the resulting intensity on the screen is so badly affected by high-frequency speckle noise that the QR code cannot be read (Fig. 1(b)). However, when two or more patterns can be used, the patterns result in complementary intensity patterns on the distal end in such a way that the resulting time-averaged intensity closely resembles the target image. Increasing the number of patterns per frame by one more than doubles the structural similarity index (SSIM) [17] with the target image. Increasing the number to five patterns leads to a very smooth background intensity. The source code for this simulation, including all simulation parameters, is available in [Code 1](#) [18].

When an image affected by speckle is projected multiple times, the speckle artifacts will average out eventually, simply because the location of the speckles does not overlap. However, by a suitable design of our algorithm, the details of which will be discussed in the next section, our image projection quality improves much more rapidly than this averaging effect. This is visible in the cross-correlation of the individual intensities corresponding to the individual patterns. After subtracting the target intensity, these patterns anti-correlate, indicating that they complement each other (Fig. 1(c)). For two patterns, the normalized cross-correlation is  $-0.89$ , and for five patterns it is around  $-0.25$ , roughly scaling with  $-1/(N-1)$ , with  $N$  being the number of patterns per target intensity. Had the patterns been independent, this cross-correlation would average out to zero. Therefore, this is an indication that the algorithm finds a set of patterns that really complement each other and that our method outperforms the statistical averaging of uncorrelated speckle patterns.

Traditionally, amplitude shaping with a phase-only modulator can only be done by discarding light. To this end, an aperture is added in the Fourier plane of the spatial light modulator. To change the intensity of the beam, light that is not desired for the current scene can then be discarded by directing it outside of the area of the aperture [19–23]. The maximum amount of power that can be transmitted is subject to two constraints: (i) the amount of light that is incident on the wavefront shaper; (ii) the spatial structure of the light that is required on the wavefront shaper. As the wavefront shaper can only attenuate the beam, it cannot freely redistribute the light in such a way that all the power is used efficiently. Therefore, when the illumination intensity on the wavefront shaper overlaps with the desired intensity, light can be excited with high fidelity and little loss of signal. When the required intensity and the illumination do not overlap, a loss in fidelity and a loss in power will be observed [24]. TAP circumvents this limitation by employing complementary speckle patterns. Therefore, most of the incident light can be efficiently redistributed. If attenuation of the target intensity is required, some form of a spatial filter would have to be used.

A requirement for our method to be applicable is that frames can arrive in quick succession rather than exactly simultaneously. Therefore, we opted to use a DMD. DMD's can only shape binary amplitude, but using them in conjunction with a spatial filter, to which the fiber can serve itself, allows to use them as a phase shaper just like an SLM, although with a significant loss in shaping efficiency [21–23]. The advantage of using a DMD is that much higher frame rates can be obtained, which is crucial for our algorithm to succeed. Our DMD can display up to 22800 patterns per second, which is typical for most DMD's, resulting in 43 microseconds per pattern, or an effective frame rate of  $22800/N$ . Many applications do not require any faster frame rates, ranging from virtual reality ( $\approx 60$  up to 240 frames per second [25]) to lithography, single photon additive manufacturing (which is to a good approximation purely dosage dependent), and many forms of optogenetics (around 70 frames per second for most applications [26]).

As an example of a possible application of simultaneous multiple-depth image projection: a particular neuron is to be stimulated in an opto-genetic experiment while neurons behind it should not be affected by the out-of-focus light [16]. Our algorithm enables us to specify where the light should go instead.

We believe that our approach paves the way for more general and advanced methods of imaging or excitation. The algorithm does not defy the laws of physics, but explicitly tries to find the best possible solution for image projection given the limitations imposed by the optical system and the number of patterns available. Although we show the principle for a stable fiber at a single wavelength, this method can be employed for any kind of optics whose performance can be accurately modeled.

In the next section, the details of the algorithm are discussed. We then demonstrate the performance of our algorithm using a multimode fiber and we demonstrate the performance enhancement that can be obtained using one or multiple patterns. We also show - to a limited extent - that images can be projected at multiple distances from the fiber facet simultaneously.

## 2. TAP algorithm description

We formulate image projection as a minimization problem. We write a *forward model*, essentially a tiny simulation, which predicts the measured intensity for a given set of input patterns. We then compare the simulated image to the target intensity and optimize the input patterns such that the predicted intensity matches the target intensity. An overview of all the variables is given in Table 1.

**Table 1. Overview of all the variables used throughout the text along with their shapes**

Symbol	Name	Type	Columns	Rows
$M$	Number of rows/columns on the target	Scalar		
$N$	Number of patterns per frame	Scalar		
$N_{DMD}$	Number of rows/columns on the DMD	Scalar		
$n_z$	Number of z-positions	Scalar		
$d$	Observed distances from the distal fiber facet	Vector		$n_z$
$W$	Weighting function	Matrix	$M^2$	$n_z$
$\vec{A}_{DMD}$	Exit wave of the laser from the DMD facet	Vector		$N_{DMD}^2$
$\Phi$	Applied phase on the DMD [0-2 $\pi$ ]	Matrix	$N$	$N_{DMD}^2$
$\mathbf{T}$	Transmission Matrix (TM)	Matrix	$n_{prox}^2$	$n_{dist}^2$
$\Psi$	Binary patterns on the DMD	Matrix	$N$	$N_{DMD}^2$
$\vec{E}_{in}$	Incident field on TM	Vector		$n_{prox}^2$
$\vec{E}_{out}$	Exit wave TM	Vector		$n_{dist}^2$

First of all, we need a metric to compare the predicted intensity to the target image. Here, we will use Structural Similarity (SSIM) as it more closely mimics human perception than for instance a Euclidean distance norm [17]. As the spatial extent to which an image can be projected may be limited, we weigh the importance of the structural similarity with an image of the light transmission through our optical system, which we call weighted SSIM (WSSIM). We now define our optimal pattern as the pattern that maximizes the WSSIM between the projected time-averaged and the target image.

$$\arg \min_{\Psi} \mathcal{L}_{WSSIM} = \arg \min_{\Psi} \left( 1 - \text{WSSIM}(\vec{I}_t, \langle I(\Psi) \rangle, \vec{W}) \right). \quad (1)$$

Here,  $\vec{I}_t$  is the desired target image, arranged into a column vector,  $\langle I(\Psi) \rangle$  is the time-averaged intensity given by a set of patterns  $\Psi$  on the DMD, and  $\vec{W}$  is a column vector describing the

reachable area of the image projection device. Every pattern on the DMD is rearranged into a single column of  $\Psi$ . Now, in order to find the best set of patterns to project image  $\vec{I}$ , we have to solve Eq. (1) for  $\Psi$ , for which we will use a gradient-based optimization method (Adam [27]).

To enable a flexible and straightforward implementation of the gradient, we use an automatic differentiation package called JAX [28], which can also perform calculations on a graphics card (in our case an NVIDIA 2080 Ti). As WSSIM is a normalized metric, its gradient is not normalized. Therefore, in our code, we multiply the WSSIM with the number of pixels on the distal end to ensure that the gradient is stable with respect to the number of pixels.

### 2.1. Forward model

The time-averaged intensity can be written as the mean of the intensities resulting from the single frames on the DMD.

$$\langle I(\Psi) \rangle = \frac{1}{N} \sum_{i=0}^N \vec{I}(\Psi_i). \quad (2)$$

Here,  $N$  is the number of patterns per frame,  $\vec{\Psi}_i$  is the  $i$ -th column of  $\Psi$ , corresponding to a single DMD pattern, and  $\vec{I}(\Psi_i)$  is the resulting intensity of projecting the  $i$ -th DMD pattern.

The setup will be discussed later in Section 3. To derive the algorithm, it suffices to know that the system's response can be predicted for any input field  $\vec{E}_{in}$  using the following equation [29]:

$$\vec{E}_{out} = \mathbf{T} \cdot \vec{E}_{in}. \quad (3)$$

Here,  $\vec{E}_{in}$  is the input field arranged into a column vector,  $\mathbf{T}$  is the TM and  $\vec{E}_{out}$  is the resulting output field. The TM fully describes the input-output response of the optical system, and it can be measured in multiple ways [2,3,6,7,29–34]. For now, we presume it to be known.

For our setup, the intensity on the camera for a given DMD frame can be written as

$$\vec{I}(\Psi_i) = |\mathbf{T} \cdot \vec{E}_{in}(\Psi_i)|^2, \quad (4)$$

where  $\vec{E}_{in}(\Psi_i)$  is a function describing the incident electric field on the fiber induced by the  $i$ -th input pattern, and the absolute value squared is applied on every element. This also reveals that the electric field to intensity conversion makes this a non-linear problem. The exact relationship between the pattern on the wavefront shaper and the electric field incident on the TM is dependent on the exact layout of the optical system and in which basis the TM is measured.

Putting all of these equations together, we arrive at the following forward model:

$$\langle I(\Psi) \rangle = \frac{1}{N} \sum_{i=0}^N |\mathbf{T} \cdot \vec{E}_{in}(\Psi_i)|^2. \quad (5)$$

For our optimization procedure to run efficiently, the gradient of  $\mathcal{L}_{WSSIM}$  with respect to  $\Psi$  should be smooth. Unfortunately, the DMD that is employed in the setup can only perform binary amplitude modulation, meaning that it will not have a continuous gradient and hence the optimization algorithm would perform very poorly. However, just like an SLM can be used as an amplitude shaper by redirecting light outside of an aperture, a DMD can be used as a phase shaper by using it in a first diffraction order. The behavior of an SLM, which is capable of continuous phase retardation, can be appropriately modeled and does have a smooth gradient. Therefore, in our forward model, we will approximate the behavior of the DMD by modeling it as if it were an SLM. After optimizing the intensity distribution, the incident field on the proximal end of the fiber is computed based on the patterns of this simulated SLM. Based on this incident field, a Lee Hologram is computed for the DMD such that it results in a similar incident electric field [21]. For the specifications that we employed in our system, the fidelity of this procedure is

around 98%. An alternative approach would be to simulate the binary pixels with a function that ranges between 0 and 1. In our case this resulted in lower performance, the reason for this was not further investigated. As the total phase space (the number of unique patterns that can be generated) is greater for the simulated phase shaper, and for reasons of computational efficiency, we modeled the theoretical SLM to have only one-fourth of the pixels of the original DMD in both directions.

The main computational bottleneck is the computation of the dot product in [Eq. (4)]. The number of floating-point operations for this dot product scales with  $\mathcal{O}(n_{\text{dist}}^2 \log n_{\text{dist}}^2)$ , with  $n_{\text{dist}}$  the number of pixels on the distal end [35]. Therefore, ensuring that the number of elements  $T$  is kept as small as possible will improve the speed at which the algorithm can run. We measure the TM on a grid of  $45 \times 45$  pixels on the proximal end and  $256 \times 256$  pixels on the distal end, yielding a TM with roughly 132 million elements. Under specific circumstances, such as when image projection at a larger distance from the fiber facet is required, the spatial extent of the area reachable by the fiber will increase, and therefore the TM will get even larger.

To minimize the number of elements in the TM without sacrificing information, we first crop the TM on the distal end to a square extending exactly around the edges of the fiber. Then, fields on the distal end are down-sampled in the Fourier domain, as the fiber only supports a limited range of spatial frequencies due to the limited numerical aperture. This enabled us to downsample the camera frames down to  $43 \times 43$  pixels. Further enhancements could be obtained by, for instance, cropping the matrix on the proximal end, computing a low-rank approximation of the TM, or masking the entries on the distal end in areas where light exiting the fiber can not reach, but these options were not investigated.

Typically, optimizing a single frame requires around 1-5 seconds per frame, but it strongly depends on the number of patterns that are generated in parallel (as the graphics card cannot be fully utilized for just a single pattern), and on the size of the TM. SSIM is not guaranteed to be a convex metric, but in our case this did not result in a stagnation of our algorithm, although in rare cases an area with smooth intensity was skipped. An example implementation of our algorithm is available in [Code 1](#) [18], both for our example setup and an example implementation using a TM. The code to compute the structural similarity was adopted from scikit-image [36] and ported to JAX. All figures are generated using Matplotlib [37] and the final layout was performed in Inkscape.

Other algorithms, such as an adopted Gerchberg-Saxton algorithm [6], can be employed for image projection as well. In [Code 1](#) [18], we also confirmed that GS algorithm performs similar to our TAP implementation for a single pattern per frame, with the GS algorithm slightly outperforming our inverse modelling approach. We attribute the difference in performance to the automatic intensity scaling that is inherent in a GS algorithm.

### 2.1.1. Simultaneous image projection at multiple planes

Instead of projecting an image in a single plane, the algorithm can be extended to allow for multiple distances behind the fiber plane. To this end, we adopted [Eq. (4)] to allow for different propagation distances behind the fiber.

$$\langle I(\Phi, z) \rangle = \frac{1}{N} \sum_{n=0}^N \left| \mathcal{P}_z \left[ T \cdot \vec{E}_{\text{in}}(\Phi) \right] \right|^2. \quad (6)$$

Here,  $\mathcal{P}_z[\cdot]$  stands for Fresnel propagation over a distance  $z$  using the angular spectrum propagator as described for instance in Ref. [38]. The cost function can now be extended for different depths.

$$\mathcal{L}_{\text{SSIM}, z} = \sum_{i=0}^{n_z} \mathcal{L}_{\text{SSIM}}(I_i(z), \langle I(\Phi, z) \rangle, W(z)). \quad (7)$$

Here,  $I_i(z)$  is the target image at depth  $z$ , and  $W(z)$  is the reachable area at depth  $z$ .

The effective area that can be reached by the fiber is dependent on the distance from the fiber. The size of the area that can be reached expands upon propagation, but towards the edges of the fiber the effective intensity will decrease as not all the light from the fiber can be used, and the numerical aperture (NA) will decrease leading to a loss of resolution [39]. As a crude approximation, the furthest point in the radial direction that can be reached by the fiber would be  $a + NAz$ , with  $a$  the fiber radius and  $z$  the propagation distance from the facet. This corresponds to a beamlet exiting the fiber from the edge in a radial direction outwards, at an angle limited by the NA of the fiber. In practice, the area in which images can be projected expands at about half that speed, as more beamlets are necessary to render a scene. Therefore we scale the target images by a factor  $\frac{a+NAz/2}{a}$  to ensure that the image details expand at approximately the same rate as the light leaving the fiber would.

This type of multi-plane image projection is subject to two constraints. First of all, Fresnel propagation preserves the total amount of power, so it is impossible to render two scenes at different depths with a different brightness level without dumping the excessive light somewhere in between. Therefore, we equalize the averaged weighted brightness in all scenes. Moreover, the frames cannot be arbitrarily close to each other. The minimum distance requirement is dependent on the similarity of the target intensities and on the depth of field of the fiber, in our case around fifteen  $\mu\text{m}$ . From a ray optics perspective, this can be understood in the following way: if a particular area is bright in the first plane, but dark in the second plane, the distance between the planes must be sufficient for light to move away, where the maximum angle under which it can travel is dominated by the numerical aperture of the fiber. The exact distance is dependent on the details of the intensity distribution of the target images in both planes.

To ensure that the field of view is sufficiently large for the entire area in which we intend to project an image, the fields on the distal end of the fiber are padded in real space and Fourier space on the distal end. The padding is done such that the field of view safely incorporates the entire area that can be reached by the fiber over a distance of up to  $100\mu\text{m}$  from the distal fiber facet. This may sound counter-intuitive as the TM was first cropped to reduce the number of elements, but as this operation only has to be done for the fields exiting the fiber, the algorithm is in fact much more efficient.

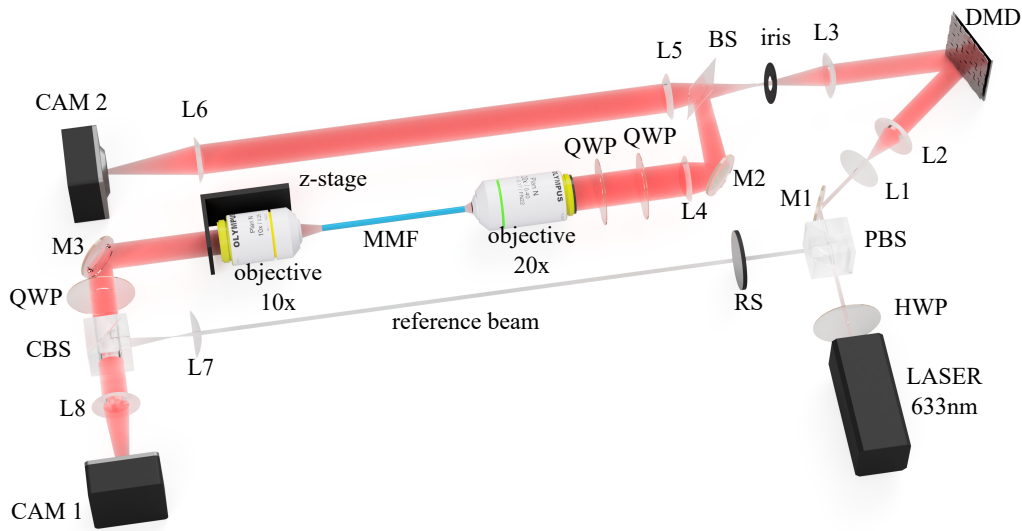
For optimal performance, it is important to scale the target brightness such that a significant portion of the light available to the DMD is also used for image projection. When the requested brightness requires a total power that is higher than what the system can provide, it will lead to artifacts in the image projection. Similarly, when the desired brightness of the target is much lower than the achievable brightness, a lot of light will have to be scattered away and this will negatively affect the dynamic range. To this end, we normalize the illumination on the DMD and we also scale values in the TM by a constant such that for an incident field that is supported by the fiber, the transmitted power is around one. As the last step, before we start the optimization procedure, we run an additional optimization procedure that just optimizes the patterns in such a way that the total power transmitted through the optical system is maximized. The resulting transmitted power is then used to normalize the target images. This last step ensures that light coming from the DMD is efficiently used.

In addition to normalizing the values of the intensity projection, we have also to normalize the gradient, since this is driving the optimization algorithm. In our case, the gradient corresponds to the change in radians per pixel for every iteration. Therefore, a typical update step should be no more than  $2\pi$ , as otherwise only wrapping would occur. To this end, we scale the gradient in the first 10% of the iterations, such that the mean squared gradient is 0.1 radians, by multiplying the gradient with a scalar. This scaling is then kept at the same value, but as the algorithm gets closer to a solution, the size of the gradient will decrease and therefore the scaling is no longer required. We initialize  $\Phi$  as the phases required to generate a smooth beam coming out of the

fiber, mainly with the intent of avoiding starting out with optical vortices in the fields leaving the fiber, which can be difficult to remove [11].

### 3. Experimental setup

To demonstrate our algorithm, we project images on the distal end of a multimode optical fiber. The setup is shown in Fig. 2. The DMD is placed in the Fourier plane of the input facet of the optical fiber, and the light leaving the optical fiber is measured on CAM 1. The electric field for every input point is measured by off-axis phase-shifting interferometry using an external reference. The TM is measured in a basis of diffraction-limited points on the input facet of the fiber, corresponding to gratings with varying pitch centered around one-fourth of the Nyquist frequency [6,7,40]. To project a diffraction-limited point, the flatness of the DMD has to be compensated for. The amplitude of the incident light ( $\vec{A}_{\text{DMD}}$ ) and the flatness of the DMD are calibrated beforehand with the subdomain approach [6,41] using  $48 \times 48$  tiles. A post-processing algorithm then centers and downsamples the TM. In image-projection mode, the reference beam is blocked using a removable shutter.



**Fig. 2.** Optical setup. The laser beam is expanded (L1 (19mm) – L2 (500mm)) onto the DMD (DMD V-7001 SuperSpeed V-module, Vialux). The Fourier plane of the DMD is imaged (L3 (200mm) – BS – M2 – L4 (80mm) – M1) onto the proximal facet of the fiber (Step-index fiber, core diameter 50µm, numerical aperture 0.22, Thorlabs FG050LGA). On the distal end, a reference beam is mixed under an angle with the signal beam exiting the fiber. Half waveplates (Thorlabs WPH10ME-633) and quarter waveplates (Thorlabs WPQ10M-633) are used to ensure that circular polarized light is excited and analyzed, indicated as HWP and QWP, respectively. A removable shutter, RS, is used to block the reference beam. To measure the flatness of the DMD, the Fourier plane of the DMD is magnified and imaged onto CAM 2.

The incident light on the TM can now be modeled as

$$\vec{E}_{\text{in}}(\Psi_i) = S(\mathcal{F}(\vec{A}_{\text{DMD}} \odot \Psi_i)). \quad (8)$$

Here,  $\vec{A}_{\text{DMD}}$  is column vector with the (complex) amplitude incident on the DMD, which is multiplied pointwise ( $\odot$ ) with the input patterns for every input pattern. The Fourier transform is

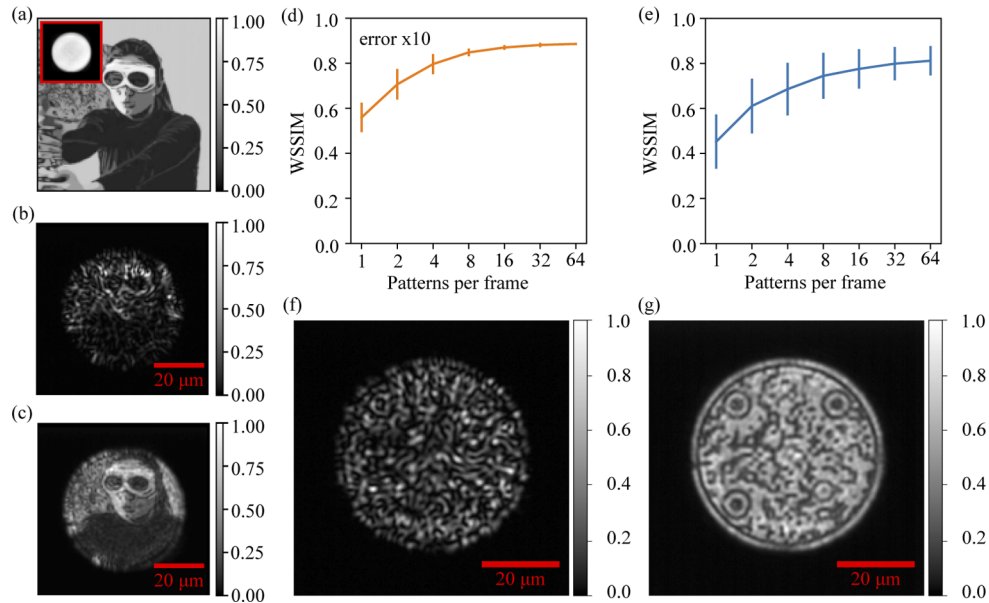


indicated by  $\mathcal{F}(\cdot)$ , and  $S(\cdot)$  is a function extracting the part of the Fourier space in which the TM was measured, and rearranging it into a column vector. The incident light on the fiber of our modeled SLM can now be written as

$$\vec{E}_{\text{in}}(\Phi_i) = S\left(\mathcal{F}\left(\vec{A}_{\text{DMD}} \odot \exp(i\Phi_i)\right)\right). \quad (9)$$

For our proof-of-concept, we measured the intensity resulting of every individual pattern to avoid synchronisation issues between the camera and the DMD. The obtained images were then averaged in time to calculate the time-averaged intensity.

As a first demonstration, we show a frame of a movie produced by Leibniz-IPHT in Fig. 3(a), which we will refer to as Movie 1, available as [Visualization 1](#). This is a rather difficult scene to project in our optical system as it has relatively large areas with uniform intensities and sharp edges between the areas. When only a single pattern were to be used, a few low-order fiber modes would have to be excited with high fidelity. This is hard to achieve in our system as it requires a high intensity from a small area of the DMD, and therefore a lot of light would be lost. This is confirmed in our experiment (Fig. 3(b)), where our algorithm fails to produce an image with sufficient quality. Using multiple patterns per frame (Fig. 3(c)) a considerable improvement in projection fidelity is seen.

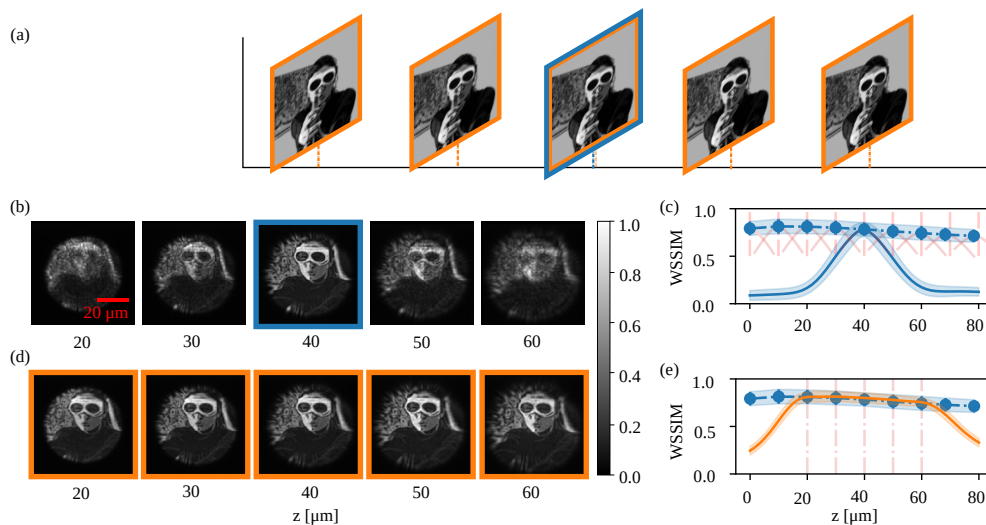


**Fig. 3.** Intensity projection using multiple patterns. A movie is projected on the distal end of the fiber. (a) Target intensity of the first imaging frame. Inset in top left: weighting function. (b) Measured intensity distribution optimized using one pattern per movie frame. (c) Measured intensity distribution optimized using 32 patterns per frame. (d) Image projection fidelity for the first frame for various numbers of patterns per frame. The algorithm was repeated for different initial states of the algorithm, and the error bars correspond to the 10 times the standard deviation in the resulting image projection quality. (e) Image projection quality averaged over the entire movie. As different frames can be more or less challenging to render, the errorbar is larger. The original image is a movie still from a video about COVID-19 measures in the research institute [42]. (f) QR code generated using a single pattern per frame. (g): QR code generated using 128 patterns per frame.

We analyze the WSSIM of the resulting intensity images. Due to resampling, the target images do not have the same number of pixels as the images that are recorded. To ensure that we do not

artificially improve our results, the recorded images are down-sampled to the target image size, rather than up-sampling the target image. This procedure can only lead to a reduction in image contrast and hence a lowering of the SSIM.

The full movie for all patterns is available in [Visualization 2](#). Rendering the first frame multiple times with different initial conditions shows that the final image projection quality is stable, especially when using a larger number of patterns per frame. For the first frame, the WSSIM increases from  $0.56 \pm 0.005$  for a single pattern to  $0.885 \pm 0.0005$  when using 64 patterns. The individual frames are visible in [Visualization 3](#). The image projection quality is dependent on the scene, though, and analyzing the entire movie for different scenes reveals that not all frames can be projected with the same fidelity (Fig. 3(e)), resulting in a larger error bar. Using this technique, it is possible to render a circular QR code on the distal end of the fiber with sufficient quality using 128 patterns per frame, but not when using a single pattern. The QR code is a very challenging target for our application, and we used a circular pattern inside the QR code to aid in the projection, as it resembles natural speckle more closely. While it was possible to render a traditional square QR code, it could only be read on very few devices. The individual frames adding up to the QR code are available as [Visualization 4](#). The QR code corresponds to a link to [Code 1](#). We verified that the QR code can be read on a Moto G2 smartphone with the free Android app "QR scanner" for Android [43], version 2.6.6.



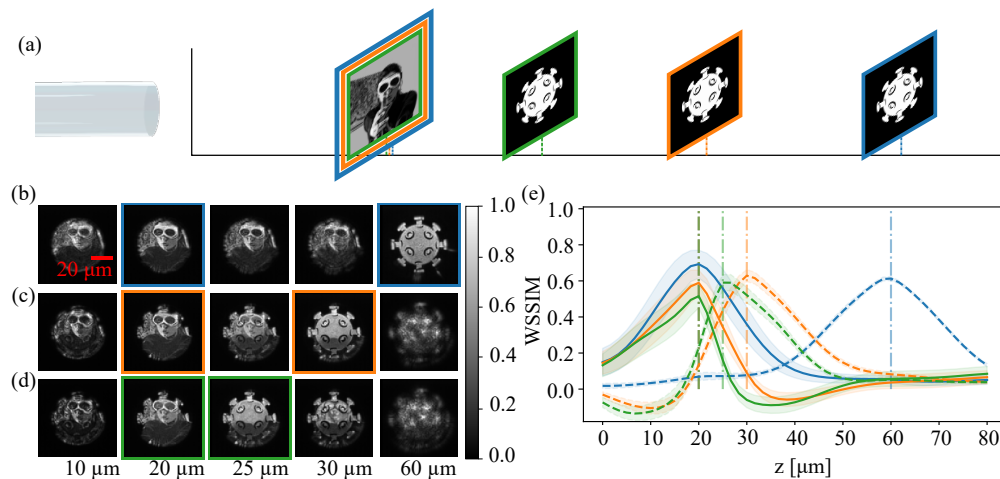
**Fig. 4.** Depth-extended image projection. (a) The intensity is either optimized for single-depth image projection at  $40\mu\text{m}$  from the fiber facet (blue outline), or for multiple depth planes ranging from  $20$  to  $60\mu\text{m}$  (orange). (b) Measured intensity for a single target depth. The image is sharp in a region of around  $15\mu\text{m}$  around the targeted imaging depth, but outside of this area, a speckle background becomes visible. (c) WSSIM for various imaging depths. The same experiment is repeated at different target imaging depths and shown in pink. The maximum performance is shown as a thick blue line and can be seen to go down due to a loss of high-frequency features. Image scale is the same as in (c). (d) Measured intensity for multiple target depths. The image remains in focus over the entire targeted imaging range of  $20$ - $60\mu\text{m}$ . (e) WSSIM analysis of the extended depth of focus. As a comparison, the maximum performance at every single depth is repeated from (c). Extending the depth of focus does not come at an expense of the attainable resolution for this target.

Using the angular spectrum method, we move the projection plane to  $40\mu\text{m}$  from the fiber facet through post-processing. The projected image is now sharp in a range of about  $15\mu\text{m}$  around the

image projection plane. Out of focus, the speckled nature of the individual patterns becomes visible. The WSSIM ranges from 0.09 out of focus to 0.8 at the target projection depth, with a full-width at half-maximum (FWHM) of about  $25\mu\text{m}$ . Repeating this experiment for different distances from the fiber facet reveals that the optimum performance goes down upon propagation. This is to be expected as light has to be spread over a larger area and the NA will go down towards the edges of the field of view [39].

Using more than one pattern per frame allows for much greater flexibility than simply improving the image projection quality at a single projection depth. As a first example, we will extend the depth of focus over a larger area. This is difficult to achieve using a single pattern, as it would require exciting some form of a propagation-invariant beam, which may not exist for the desired intensity image, or not available to this specific optical system. Contrary, when the target image can be split up into multiple individual intensity patterns, multiple beamlets can generate the desired emergent effect, although none of the individual patterns are required to have the prescribed behavior.

To extend the depth of focus, we specify that the same target should be displayed in five subsequent planes separated by  $10\mu\text{m}$ , schematically indicated in Fig. 4(a). The resulting intensity profile is essentially propagation invariant over the entire imaging range (see also [Visualization 5](#)). The FWHM is now extended to  $73\mu\text{m}$ . This range can likely be extended even more, but this has not been experimentally verified. Moreover, the averaged image projection quality does not seem to suffer from this additional requirement, which is surprising. We only see a dark area towards the right of the target image, which should be uniformly bright. This is an artifact that occasionally shows up in challenging scenes and is probably a limitation of the chosen cost function.



**Fig. 5.** Multi-depth image projection. (a) Desired target intensity consists of Movie 1 at  $20\mu\text{m}$  from the fiber facet and Movie 2 at a distance of  $60\mu\text{m}$  (blue),  $30\mu\text{m}$  (orange), and  $25\mu\text{m}$  (green). (b) WSSIM of the resulting solution with respect to Movie 1 (solid line) and Movie 2 (dashed line). Colors correspond to the different imaging depths, and a vertical line indicates the location of the image projection. The error bars correspond to the standard deviation of the entire movie.

Instead of projecting the same intensity with an extended depth of focus, we can also project independent scenes at different depths using the same set of patterns. This may be useful to 'hide' a particular message behind a different message, or in optogenetics, to avoid a certain area in the tissue, such as a neuron that is not intended to be stimulated. Furthermore, by specifying different

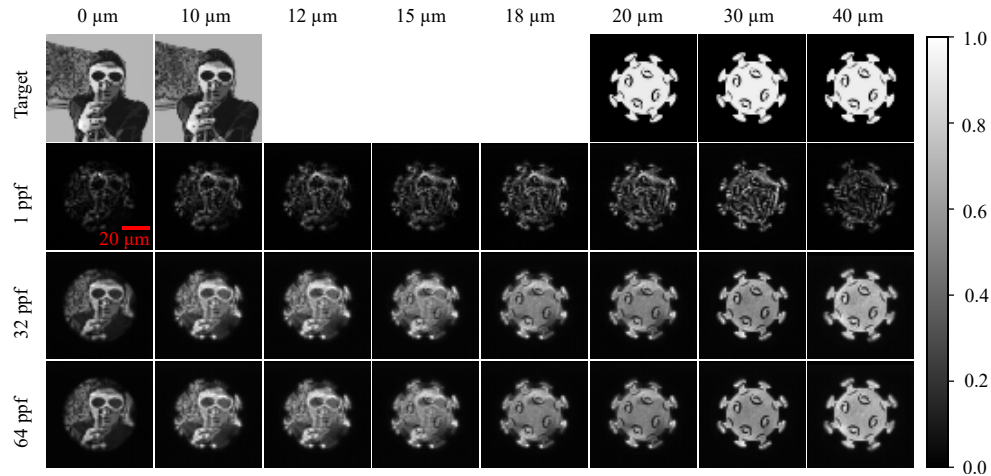
target intensity distributions at different depths, it may be possible to tailor the out-of-focus light in a particular way. For instance, the width of the WSSIM curve may be tuned, and instead of an extended depth of focus, an image with a deliberately small depth of field may be obtained.

To this end, we display two movies simultaneously, see Fig. 5. In the first plane, we display Movie 1. In the second plane, we display a rotating cartoon picture of a virus, which we'll refer to as Movie 2, available as Visualization 6. To ensure that the two intensity planes do not interfere with each other, a spacing of  $40\mu\text{m}$  between the planes is kept, which is more than twice the depth of field. The two planes can be rendered independently and no shadow from one plane is visible on the other plane. The image projection fidelity of the first plane is reduced from around  $0.85 \pm 0.05$  for single-target image projection to  $0.7 \pm 0.1$ .

When the planes are moved closer together, a shadow of the first image becomes visible in the second image, and hence the maximum WSSIM goes down. This is to be expected, as the distance between the planes is lowered to less than the depth of field of the fiber. However, the transition area between the planes becomes much sharper, which can be derived from the steepness of the curves in Fig. 5(c), and is visible in Visualization 7, right column. The WSSIM curves are no longer symmetric and the width of the peaks is reduced. Hence, we conclude that we can indeed make the depth of field both artificially deep and shallow, although the latter only at the expense of a drop in image projection quality.

As a generalization of two-plane image projection, we can project two independent target intensity distributions, both with an extended depth of field. As a loss function, we specify that Movie 1 is to be shown at 0 and  $10\mu\text{m}$  from the facet, and Movie 2 is to be shown at 20, 30 and  $40\mu\text{m}$ , see Fig. 6. This combination of different intensities at multiple planes is particularly ill-suited for single-pattern-based image projection, because upon propagation, amplitude and phase of light mix. This feature has for instance been used in the development of lensless microscopes [44–48], to reconstruct an image based on the evolution of the diffraction patterns. When this intensity is constrained in more than two planes, there need not be any electric field that will satisfy the required intensities in all planes. However, when more than one pattern can be employed, more flexibility is possible due to emergent effects.

Our experiments confirm that, when using a single pattern, severe aberrations to the target intensity distributions are visible. Using 32 or 64 patterns results in a small transition area between the scenes (see also Visualization 8). Analogously, up to five subsequent frames of Movie 1 can be displayed at different depths, with a spacing of  $10\mu\text{m}$ , which we demonstrate in Visualization 9 and which is discussed in Appendix A. Here, the image quality suffers as the scene can be too complex, but the result is surprisingly accurate. Stitching back the individual frames into a movie (Visualization 10) reveals relatively minor amounts of crosstalk.



**Fig. 6.** Double extended depth-of-field image projection. At the fiber facet and at a distance of  $10\mu\text{m}$  from the facet, we project Movie 1. Simultaneously, at  $20, 30,$  and  $40\mu\text{m}$  from the facet, Movie 2 is shown. Although some shadow of Movie 2 is visible in the first planes, the projected images are still of very reasonable quality. In between the planes, a sharp transition area is visible, shown at  $12, 15,$  and  $18\mu\text{m}$  from the facet. Scale bar is the same for all images.

#### 4. Conclusion

We have demonstrated that time-averaged image projection (TAP) can yield unprecedented image projection flexibility by decomposing a challenging intensity scene into a set of patterns that are displayed onto a wavefront modulator that is well-matched to the optical system at hand. The optical system has to be known in advance, in particular a function mapping the input-output relationship has to be available. It is not required to know exactly which kind of intensity distributions pose a challenge for the optical system, and the algorithm does not rely on any matrix inversion technique. The number of patterns required is usually considerably smaller than the number of points in the image, but more than one.

TAP provides a robust algorithm for image projection through any optical system whose performance can be modeled accurately. The algorithm provides a repeatable performance even for different initial conditions. An implementation of the algorithm for a setup using a multimode optical fiber demonstrates that a weighted structural similarity as high as  $0.885 \pm 0.0005$  can be obtained for the first frame of testing Movie 1, allowing to project a QR code with a light background. We have also shown that the behavior of the out-of-focus light can be tailored in various ways. One possibility is to extend the depth range at which the image is sharp from about  $25\mu\text{m}$  to at least  $73\mu\text{m}$ , with no detectable loss in image projection quality. Another possibility is to project two independent scenes at various depths, provided that enough spacing is present between the different planes. Lastly, we can create a sharp transition in between the two imaging planes, enabling us to generate a sharp transition from one depth to the other, at the expense of a loss of image projection quality. Lastly, it is possible to combine the two approaches and create a set of patterns that result in two independent scenes being displayed behind each other, both with an extended depth of focus, or to project slowly varying scenes in different depths using the same set of patterns.

From an application perspective, the main bottleneck for our application is the calculation of the frames to be projected on the DMD, which can take up to a few seconds per target intensity,

depending on the size of the TM. It is possible that this can be sped up using a more tailored optimization algorithm. Once this is done, the frames can be displayed with the maximum framerate available, in our case resulting in a framerate of 22800/N. For most applications, the first ten frames result in the bulk of the enhancement, resulting in a framerate of about 2280 frames per second, but multi-plane image projection tends to require more planes. Further enhancements of the image projection speed could be obtained either using only a subset of the DMD surface area allowing for a speed increase of the raw frames per second, or by reducing the number of frames required by a suitable design of the cost function, but this is dependent on the application.

In this paper, we only consider changing the projection depth of the system, but the algorithm is much more widely applicable. Instead of tuning the behavior of the light upon propagation, the behavior at different wavelengths could be tuned or even a specific change to the image when the fiber is deformed could be specified. Furthermore, multi-wavelength image projection can be attempted or 'bend-invariant' image projection by sampling transmission matrices taken for different bends of an optical fiber, and specifying that the same intensity should be obtained. As the technique is not limited to image projection through optical fibers, future applications may also involve automated alignment or targeted light delivery in applications such as optogenetics or lithography. In addition, by appropriately changing the cost function, a specific area in 3D space can be avoided, with the light free to go anywhere else, which was successfully simulated but not experimentally verified.

### A. Additional depth-multiplexed image projection

In [Visualization 9](#), we display Movie 1, but displaying every group of five patterns simultaneously in depth. The first five frames are displayed simultaneously with a spacing of 10 $\mu$ m, then the next set of five frames is displayed in the same configuration, using 32 patterns per frame. Most of these frames are similar except for the moving parts in the image. Although some artifacts are present, the desired behavior in depth is visible. Especially in frames 10-15, the hand of the main character is moving although the rest of the frame is stable. After recording the entire movie, we can display the frames in their original order. This is shown in [Visualization 10](#), rendering the original movie.

**Funding.** European Research Council (724530); Ministerstvo Školství, Mládeže a Tělovýchovy (CZ.02.1.01/0.0/0.0/15\_003/0000476); European Regional Development Fund (CZ.02.1.01/0.0/0.0/15\_003/0000476); Freistaat Thüringen (2018-FGI-0022, 2020-FGI-0032); Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale Gesellschaft; Thüringer Aufbaubank; Bundesministerium für Bildung und Forschung.

**Acknowledgments.** We'd like to thank Beatriz Silveira for a very thorough reading of the draft, and Angel Cifuentes for turning on the setup remotely.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** The transmission matrices that are specific to this optical setup are not publicly available at this time, but can be obtained upon request.

### References

1. A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," *Nat. Photonics* **6**(5), 283–292 (2012).
2. I. M. Vellekoop, A. Lagendijk, and A. P. Mosk, "Exploiting disorder for perfect focusing," *Nat. Photonics* **4**(5), 320–322 (2010).
3. T. R. Hillman, T. Yamauchi, W. Choi, R. R. Dasari, M. S. Feld, Y. Park, and Z. Yaqoob, "Digital Optical Phase Conjugation for Delivering Two-Dimensional Images through Turbid Media," *Sci. Rep.* **3**(1), 1909 (2013).
4. S. Bianchi and R. D. Leonardo, "A multi-mode fiber probe for holographic micromanipulation and microscopy," *Lab Chip* **12**(3), 635–639 (2012).
5. R. D. Leonardo and S. Bianchi, "Hologram transmission through multi-mode optical fibers," *Opt. Express* **19**(1), 247–254 (2011).
6. T. Čížmár and K. Dholakia, "Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics," *Opt. Express* **19**(20), 18871–18884 (2011).

7. J. Yoon, K. Lee, J. Park, and Y. Park, "Measuring optical transmission matrices by wavefront shaping," *Opt. Express* **23**(8), 10158–10167 (2015).
8. D. Andreoli, G. Volpe, S. Popoff, O. Katz, S. Grésillon, and S. Gigan, "Deterministic control of broadband light through a multiply scattering medium via the multispectral transmission matrix," *Sci. Rep.* **5**(1), 10347 (2015).
9. E. Tajahuerce, V. Durán, P. Clemente, E. Irlés, F. Soldevila, P. Andrés, and J. Lancis, "Image transmission through dynamic scattering media by single-pixel photodetection," *Opt. Express* **22**(14), 16945–16955 (2014).
10. S. Popoff, G. Lerosey, M. Fink, A. C. Boccarda, and S. Gigan, "Image transmission through an opaque material," *Nat. Commun.* **1**(1), 81–88 (2010).
11. M. Plöschner and T. Čížmár, "Compact Multimode Fiber Beam-Shaping System Based on GPU Accelerated Digital Holography," *Opt. Lett.* **40**(2), 197–200 (2015).
12. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode Optical Fiber Transmission with a Deep Learning Network," *Light: Sci. Appl.* **7**(1), 69 (2018).
13. B. Rahmani, D. Loterie, E. Kakkava, N. Borhani, U. Tegin, D. Psaltis, and C. Moser, "Actor neural networks for the robust control of partially measured nonlinear systems showcased for image propagation through diffuse media," *Nat. Mach. Intell.* **2**(7), 403–410 (2020).
14. M. Plöschner, B. Straka, K. Dholakia, and T. Čížmár, "GPU accelerated toolbox for real-time beam-shaping in multimode fibres," *Opt. Express* **22**(3), 2933–2947 (2014).
15. C. Maurer, A. Jesacher, S. Bernet, and M. Ritsch-Marte, "What spatial light modulators can do for optical microscopy," *Laser Photonics Rev.* **5**(1), 81–101 (2011).
16. E. Ronzitti, C. Ventalon, M. Canepari, B. C. Forget, E. Papagiakoumou, and V. Emiliani, "Recent advances in patterned photostimulation for optogenetics," *J. Opt.* **19**(11), 113001 (2017).
17. Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures," *IEEE Signal Process. Mag.* **26**(1), 98–117 (2009).
18. D. Boonzajer, "The time-averaged projection supporting code," GitLab (2021) [accessed 9 August 2021], <https://gitlab.com/dboonz/time-averaged-projection-supporting-code>.
19. A. Jesacher, S. Bernet, and M. Ritsch-Marte, "Colour hologram projection with an SLM by exploiting its full phase modulation range," *Opt. Express* **22**(17), 20530–20541 (2014).
20. G. Thalhammer, R. W. Bowman, G. D. Love, M. J. Padgett, and M. Ritsch-Marte, "Speeding up liquid crystal SLMs using overdrive with phase change reduction," *Opt. Express* **21**(2), 1779–1797 (2013).
21. Y.-X. Ren, R.-D. Lu, and L. Gong, "Tailoring light with a digital micromirror device," *Ann. Phys.* **527**(7-8), 447–470 (2015).
22. S. Turtaev, I. T. Leite, T. Altwegg-Boussac, J. M. P. Pagan, N. L. Rochefort, and T. Čížmár, "High-fidelity multimode fibre-based endoscopy for deep brain in vivo imaging," *Light: Sci. Appl.* **7**(1), 92 (2018).
23. S. A. Goorden, J. Bertolotti, and A. P. Mosk, "Superpixel-based spatial amplitude and phase modulation using a digital micromirror device," *Opt. Express* **22**(15), 17999–18009 (2014).
24. I. T. Leite, S. Turtaev, D. E. Boonzajer Flaes, and T. Čížmár, "Observing Distant Objects with a Multimode Fiber-Based Holographic Endoscope," *APL Photonics* **6**(3), 036112 (2021). Publisher: American Institute of Physics.
25. Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, and S. Yoshimura, "A psychophysical study of improvements in motion-image quality by using high frame rates," *J. Soc. Inf. Disp.* **15**(1), 61 (2007).
26. A. Malyshev, R. Goz, J. J. LoTurco, and M. Volgushev, "Advantages and Limitations of the Use of Optogenetic Approach in Studying Fast-Scale Spike Encoding," *PLoS One* **10**(4), e0122286 (2015). Publisher: Public Library of Science.
27. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," ArXiv (2017). \_eprint: 1412.6980.
28. J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne, *JAX: composable transformations of Python+NumPy programs* ([github.com](https://github.com/google/jax), 2018). Publication Title: [github.com](https://github.com).
29. I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**(16), 2309–2311 (2007).
30. S. Li, C. Saunders, D. J. Lum, J. Murray-Bruce, V. K. Goyal, T. Čížmár, and D. B. Phillips, "Compressively sampling the optical transmission matrix of a multimode fibre," *Light: Sci. Appl.* **10**(1), 88 (2021).
31. P. Pai, J. Bosch, and A. P. Mosk, "Optical transmission matrix measurement sampled on a dense hexagonal lattice," *OSA Continuum* **3**(3), 637–648 (2020).
32. G. S. D. Gordon, M. Gataric, A. G. C. P. Ramos, R. Mouthaan, C. Williams, J. Yoon, T. D. Wilkinson, and S. E. Bohndiek, "Characterizing Optical Fiber Transmission Matrices Using Metasurface Reflector Stacks for Lensless Imaging without Distal Access," *Phys. Rev. X* **9**(4), 041050 (2019).
33. A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques," *Opt. Express* **23**(9), 11898 (2015).
34. E. G. van Putten and A. P. Mosk, "The information age in optics: Measuring the transmission matrix," *Physics* **3**, 22 (2010).
35. D. Coppersmith, "Rapid Multiplication of Rectangular Matrices," *SIAM J. Comput.* **11**(3), 467–471 (1982).
36. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, t. s.-i. contributors, "scikit-image: image processing in Python," *PeerJ* **2**, e453 (2014).
37. J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Comput. Sci. Eng.* **9**(3), 90–95 (2007).

38. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill Physical and Quantum Electronics Series (Roberts & Company, 2005).
39. T. Čižmár and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," *Nat. Commun.* **3**(1), 1027 (2012).
40. M. Kim, W. Choi, Y. Choi, C. Yoon, and W. Choi, "Transmission matrix of a scattering medium and its applications in biophotonics," *Opt. Express* **23**(10), 12648–12668 (2015).
41. T. Čižmár, M. Mazilu, and K. Dholakia, "In situ wavefront correction and its application to micromanipulation," *Nat. Photonics* **4**(6), 388–394 (2010).
42. D. Siegesmund, "Lasergirl jagt den Corona Virus," (2020). <https://www.youtube.com/watch?v=Nc3KwC1GzrM>.
43. Sevegame, "QR Scanner – Apps on Google Play, version 2.6.6," (2021). <https://play.google.com/store/apps/details?id=com.kitkats.qrscanner>.
44. D. W. Noom, D. E. Boonzajer Flaes, E. Labordus, K. S. Eikema, and S. Witte, "High-speed multi-wavelength Fresnel diffraction imaging," *Opt. Express* **22**(25), 30504–30511 (2014).
45. R. Saxton and W. Gerchberg, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *Optik* **35**, 237–246 (1972).
46. L. J. Allen and M. P. Oxley, "Phase retrieval from series of images obtained by defocus variation," *Opt. Commun.* **199**(1-4), 65–75 (2001).
47. L. Loetgering, R. Hammoud, L. Juschkin, and T. Wilhein, "A phase retrieval algorithm based on three-dimensionally translated diffraction patterns," *EPL* **111**(6), 64002 (2015).
48. W. Bishara, T.-W. Su, A. F. Coskun, and A. Ozcan, "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution," *Opt. Express* **18**(11), 11181–11191 (2010).