Dynamic noninvasive imaging through turbid media under low signal-noise-ratio

Yuan Yuan¹,² and Hui Chen¹,²,³

¹ Electronic Material Research Laboratory, Key Laboratory of the Ministry of Education and International Centre for Dielectric Research, 710049, Xi’an Jiaotong University, People’s Republic of China
² These two authors contributed equally.
³ Author to whom any correspondence should be addressed.

E-mail: chenhui@xjtu.edu.cn

Keywords: non-line-of-sight, non-invasive imaging, scattering media, dynamic imaging, optical transfer function, second-order correlation, phase retrieval

Abstract

In turbid media, scattering of light scrambles information of the incident beam and represents an obstacle to optical imaging. Noninvasive imaging through opaque layers is especially challenging for reliable image reconstruction and dynamic objects. We here propose a solution to these problems: rather than using the full point-spread function or its Fourier transform (optical transfer function, OTF), the wave distortions in scattering layers can be characterized and diffraction-limited imaging performed using only the phase of OTF. Based on this understanding, we develop a method that exploits the redundant information from multiple measurements, which reliably yields OTF phases within several iterations. This method enables noninvasive imaging through turbid media with low signal-to-noise ratios in the measurements, which is not possible with previous methods. We then demonstrate noninvasive video imaging of a moving object hidden between scattering layers at 25–200 Hz. This imaging approach may inspire many other applications in scattering materials.

1. Introduction

The ability to look through light-scattering media is important in various fields ranging from life sciences [1] to nano-technology [2]. Among all the means, imaging with visible light offers better resolution than microwave and ultrasound imaging and is still considered safe for samples in most scenarios, such as biomedical tissues. In optically-thick random media, wavefront distortions due to multiple light scattering are very complicated, so the wavefront correction methods in conventional adaptive optics [3] provide limited improvement [4], which only compensate for low spatial frequency perturbations such as atmospheric turbulence and aberrations induced in translucent tissues [5]. One successful approach in strongly scattering medium is to selectively detect ballistic photons that travel through the medium in a straight line [6–8]. Unfortunately, its working depth is as shallow as a few mean free paths (MFPs) because the ballistic light flux is exponentially extinguished as the depth increases. Photons that are scattered only once [9] or slightly so that the overall propagation direction is maintained, such as photo-acoustic microscopy [10–12], can also be exploited to image target objects located deeper than ten MFPs with limited resolution.

There are also numbers of other innovative strategies which use scattered waves in full. Pioneering work [4, 13–20] utilizing the spatial degrees of freedom of light have been demonstrated including wavefront shaping [15–17] and phase conjugation [18–20], which are usually combined with a remarkable speckle correlation phenomenon known as the ‘memory effect’ [21, 22] for imaging behind tens of micron- to millimeter-thick tissues. These techniques require the presence of a guide-star in the object plane and the
use of a spatial light modulator or a digital mirror device to generate the needed phase map. Recently, noninvasive optical imaging strategies which remove those requirements have been proposed and demonstrated [23, 24], which are similar to the speckle interferometry first proposed by Labeyrie [25] in 1970, which aims to yield the diffraction-limited autocorrelation of astronomical objects in spite of atmospheric turbulence. These strategies exploit the speckle correlations caused by the ’memory effect’ for scattering and use phase retrieval for image reconstruction. However, in these methods, the image-retrieval process is undependable and time-consuming with stagnation problems, which deteriorates with low signal-to-noise ratios (SNRs) in realistic measurements. This unreliability becomes a major bottleneck for looking through scattering media. It is still a big challenge not only for imaging in weak light environments, but also for dynamic imaging. Thus, present noninvasive methods are still far from a realistic application.

To overcome the above difficulties, we propose that, when a strong scattering system contains sufficient random scatterers to satisfy the ergodic-like condition [26], the wave distortions can be effectively characterized in the Fourier frequency domain using only the Fourier phase of the PSF (the optical transfer function (OTF) phase), rather than the full PSF [27]. Once the OTF phase is obtained, diffraction-limited images can be analytically deconvolved without any iterative process. Based on this perspective, we develop a framework, named multi-frame OTF retrieval engine (MORE), to non-invasively recover the OTF phase of a scattering imaging system. It exploits the redundant information from multiple measurements to address the unreliability problem in conventional methods, resulting in a highly stable and rapid convergence in the reconstruction. Even in low SNR situations which are inaccessible with previous methods, it still stably and rapidly yields unique solutions in spite of noisy data.

It was generally believed that, for a strong scattering imaging system, a full PSF is necessary for recovering diffraction-limit images [28–33]. Since the random fluctuation of a PSF will cause an error amplification during the deconvolution, an iterative process is usually required to solve this problem even after the PSF is precisely known [28, 29, 31–33]. Recently, there are works using non-iterative deconvolution in imaging through scattering: 1) the magnitude of the OTF (MTF) was artificially changed to an empirical exponential form, where the exponential coefficient varied in different experiments and were determined empirically [34]; 2) multiple PSFs caused by point sources on different axial positions were measured and summed together, which flattened the MTF and suppressed the error amplification in the deconvolution [35]. These methods tried to minimize the influence of MTF at the cost of more measurements or calibrations. Our theory is fundamentally different. We completely avoid the MTF involved in the deconvolution and prove that the phase of the OTF is enough to analytically solve the diffraction-limit images. This not only avoids an iterative process, but also reduces the difficulty in characterizing the wave distortions because only half information of the PSF is needed.

Since MORE recovers OTF phases, it enables dynamic imaging through turbid media. Moreover, MORE preserves the information of the relative positions and orientations of an object at different moments. This allows us to simply put together all the recovered images to create a video without worrying about image misalignment. We realized video imaging at 25 ~ 200 Hz for a time-varying object hidden behind turbid media. As a non-invasive approach, our method does not require any calibration or pre-process. It starts with capturing all the speckle patterns at once, and then performs post-reconstructions for the images.

It is worthwhile to point out that, if the previous imaging methods [23, 24] were used for dynamic imaging, each frame would have to be individually reconstructed with a conventional phase retrieval algorithm. In the lack of the information of the relative positions and orientations, it is very difficult to track the variance and movement of the object to make a video, in addition to the unreliability of their convergence.

2. Physical mechanism

Figure 1 depicts the schematic for non-invasive imaging objects hidden behind a diffuser. An LED with a wavelength of $\lambda = 633$ nm and a bandwidth of $\sim 15$ nm illuminates the object through the diffuser. The light travels back from the object and gets scattered by the diffuser again, generating a speckle-like intensity pattern on a camera. The diffuser is a piece of 220-grit ground glass. The distance from the object to the diffuser is $z \approx 200$ mm. The distance between the diffuser and the camera is $d \approx 300$ mm. An aperture with a diameter of $D \approx 6$ mm is place in front of the diffuser to limit an area which the scattered light passes through.

Within a ‘memory effect’ range, spherical waves from nearby points on the object suffer similar distortions and produce shifted but almost identical speckle patterns, $S(x, y)$, on the image plane, which is the PSF of the system. For that isoplanatic system with incoherent illumination, the detected intensity $I(x, y)$ on the camera is a superposition of those speckle patterns, which can be written as a convolution of the
speckles are captured by the camera's sensor to satisfy the ergodic-like condition [26, 36, 37], the MTF can be derived as (see appendix A):

\[ \tilde{M}(u, v) = \tilde{S}(u, v) \cdot \tilde{O}(u, v) \]  \hspace{1cm} (1)

where the tilde denotes the two-dimensional Fourier transform, \( u \) and \( v \) are spatial frequency coordinates, \( \tilde{S}(u, v) \) is the OTF of the system.

Rewrite equation (1) as

\[ \tilde{I}(u, v) e^{-i\Phi_3(u, v)} = |\tilde{S}(u, v)| \cdot |\tilde{O}(u, v)| e^{i\Phi_0(u, v)} \equiv \tilde{M}(u, v), \]  \hspace{1cm} (2)

where \( \Phi_3(u, v) \) is the OTF phase, \( \Phi_0(u, v) \) is the Fourier phase of the object. The MTF, \( |\tilde{S}(u, v)| \), acts as a spatial frequency filter on \( |\tilde{O}(u, v)| \).

In a scattering system with enormous numbers of random scatterers, when more than thousands of speckles are captured by the camera’s sensor to satisfy the ergodic-like condition [26, 36, 37], the MTF can be derived as (see appendix A):

\[ |\tilde{S}(u, v)| \propto \{|T \ast T|(u, v) + \delta_D\}^{1/2}, \]  \hspace{1cm} (3)

where \( \delta_D \) is a delta-like peak. \( T(\xi, \eta) \) represents the attenuation of the light intensity at position \((\xi, \eta)\) on the turbid layer, which is proportional to squared modulus of the pupil function. As shown in figure 6, with an even light intensity attenuation over the circular-shape diffuser, the MTF has a gentle slope well below the turbid layer, which is proportional to squared modulus of the pupil function. As shown in figure 6, with an even light intensity attenuation over the circular-shape diffuser, the MTF has a gentle slope well below the turbid layer, which is proportional to squared modulus of the pupil function.

To better understand \( \tilde{M}(u, v) \), we recall that, for a scattering- and aberration-free lens system, its OTF is a real function, which can be formulated in a similar form:

\[ \tilde{I}_d(u, v) = |\tilde{S}_d(u, v)| \cdot |\tilde{O}(u, v)| e^{i\Phi_0(u, v)} \equiv \tilde{M}_d(u, v). \]  \hspace{1cm} (4)

Its MTF, \( |\tilde{S}_d(u, v)| \propto T^{1/2} \ast T^{1/2} \), plays the similar role, defining the diffraction-limit image, the inverse Fourier transform of \( \tilde{M}_d(u, v) \).

Equations (2) and (4) show that, the difference between these two systems is that the scattering system has a non-even OTF’s phase, \( e^{-i\Phi_3(u, v)} \). As long as the OTF’s phase is determined, the diffraction-limit image can be directly computed:

\[ M(x, y) = \mathcal{F}^{-1}\{\tilde{I}(u, v) e^{-i\Phi_3(u, v)}\} , \]  \hspace{1cm} (5)

where \( \mathcal{F}^{-1} \) denotes the inverse Fourier transform. Since \( I(x, y) \) is a bounded real-valued function, its
Fourier transform $\tilde{I}(u,v)$ is Hermitian symmetric. Similarly, the OTF is also Hermitian symmetric and $\Phi_S(u,v)$ is odd symmetric. Therefore, $\tilde{I}(u,v)e^{-i\Phi_S(u,v)}$ is also Hermitian symmetric. Its inverse Fourier transform $M(x,y)$ is real-valued. From equation (5), using only the phase of OTF can directly compute the diffraction-limit image of the scattering system. Without the MTF involved in the deconvolution operation, the noise amplification can be avoided and no iterative process or equivalent operation is required at all. In the following, we provide an experimental proof for this theory.

In the setup of figure 1, we replaced the object by a resolution chart board (figure 1(b)). The camera-captured speckle pattern is shown in figure 2(b). The PSF of the system, $S(x,y)$, was directly measured with a point source placed on the object plane. We took the Fourier transform of the measured PSF, and obtained the OTF’s phase. Using equation (5), the image of the target is directly computed from the OTF’s phase and the captured speckle pattern without any iterative process. The result is shown in figure 2(c). The smallest part that can be clearly resolved is group 4, Element 5 of the resolution test chart, corresponding to an effective resolution of $2^{-(\text{Group}+\text{Element}−1)/6+1}$ mm $\approx$ 20 μm, in agreement with the expected diffraction-limited resolution of $\lambda/nD \approx 18 \mu m$.

Previously, in a scattering system, a full PSF was used to deconvolve images with iterative algorithms [14], such as the Richardson–Lucy (RL) algorithm [39, 40]. Here, we repeat this operation with a Matlab RL function ‘deconvlucy’. Figure 2(d) shows the result with 600 iterations of the RL algorithm. The experiment demonstrates that, the OTF phase itself contains enough information for recovering the diffraction-limited image when lacking specific noise or source information. A more detailed comparison of the image qualities obtained by different deconvolution approaches is beyond the scope of this work and will not be discussed here.

### 3. Dynamic noninvasive imaging from multiple speckle patterns

#### 3.1. The framework of MORE

Equation (5) suggests that, once the OTF phase is determined, the motion of a dynamic object can be monitored in real-time by directly deconvolving the images from the captured speckle patterns. However, in an noninvasive scenario, the PSF or the OTF phase cannot be directly measured. Fortunately, the physical mechanism also reveals that, the OTF phase can be computed by exploiting phase retrieval. Recall equation (2), we have

$$\left|\tilde{I}(u,v)\right| = \left|\tilde{M}(u,v)\right|$$

$$\Phi_S(u,v) = \Phi_I(u,v) - \Phi_O(u,v),$$

where $\Phi_I$ is the Fourier phases of the speckle pattern. Equation (6) indicates that, the Fourier phase of the object, $\Phi_O(u,v)$, can be recovered from $\left|\tilde{I}(u,v)\right|$ using a phase retrieval algorithm such as HIO or ER [41]. The OTF phase is then retrieved with equation (7).

However, the phase retrieval from one single pattern is unreliable and even fails to give a reasonable solution, especially when the SNR in the measurements decreases. As shown in the experiments, when the data is captured at 200 Hz (5 ms exposure time), the speckle patterns are corrupted with experimental noise.
noise, which results in the failure of the HIO-ER. Please see appendices D and E for the performance of the HIO-ER under different SNRs.

To overcome the above challenge, we propose a framework to recover the OTF phase with multiple patterns. This method is developed in the spirit of the local isoplanatism effects in speckle imaging, i.e., these speckle patterns are generated from different objects (or different moments of the dynamic object) but from the same OTF. Therefore, the OTF is an instinctive and robust connection for the patterns. The redundant information across multi-patterns helps to stably and rapidly retrieve the OTF phase. This method is therefore termed ‘MORE’, which implies using more camera frames (patterns), differentiating itself from the above method with a single pattern.

As shown below, the framework starts with an initial random guess for the OTF phase, and keeps updating the OTF in a few big iterative loops (indexed by \( k \)). In each big loop, the multiple frames (indexed by \( j \)) are successively used to correct the guessed OTF’ phase:

(a) Form \( \hat{O}_{k,j}(u, v) = |\hat{I}_j(u, v)|e^{i\arg[\hat{I}_j] - \arg[\hat{O}_{k,j}]} \).

(b) Form \( O_{k,j}(x, y) = F^{-1}\{\hat{O}_{k,j}(u, v)\} \);

(c) Apply realness and non-negativity constraints to \( O_{k,j}(x, y) \), yielding \( O'_{k,j}(x, y) \):

\[
O'_{k,j}(x, y) = \begin{cases} 
\text{Re}\{O_{k,j}(x, y)\}, & (x, y) \in \Gamma \cap \{\text{Re}\{O_{k,j}(x, y)\} \geq 0\} \\
0, & (x, y) \notin \Gamma \cup \{\text{Re}\{O_{k,j}(x, y)\} < 0\} 
\end{cases}
\]

where \( \text{Re}\{\} \) indicates taking a real part, and \( \Gamma \) is the support (the region of the object). See appendix B for the determination of the support.

(d) Form \( \hat{O}'_{k,j}(u, v) = F\{O'_{k,j}(x, y)\} \);

(e) Update the OTF phase: \( \arg\{\hat{S}_{k+1}(u, v)\} = \arg\{\hat{I}_j(u, v)\} - \arg\{\hat{O}'_{k}(u, v)\} \).

(f) Repeat step 1 to 5 with the next frame until all the frames are used.

(g) Repeat step 1 to 6 to update the OTF phase until the criterion of convergency is reached.

Here, the constraints in step 3 are based on the physical facts that the image must be real and non-negative. \( O \) and \( \hat{O} \) represent the calculated image and its Fourier transform. \( \arg\{\} \) is the phase operator. \( \hat{I}_j(u, v) \) is the Fourier transform of the \( j \)th speckle pattern.

3.2. Dynamic noninvasive imaging with MORE

Using an independent phase retrieval for each speckle pattern will result in an uncertain position shift of the each image [42] (see figure 9(c)). As we can see from equation (6), the retrieval process uses the Fourier magnitude of the speckle pattern, which is invariant for any position shift of the object. Different from the phase retrieval for each single object separately, MORE recovers the OTF phase that is used to deconvolve all the images (see figure 9(b)). As we can see from the following formulism, if an additional phase of \( 2\pi \cdot \mathbf{u} \cdot \mathbf{t} \) is appended to the OTF, all the reconstructed images from the same OTF phase will have the same spatial shift of \( \mathbf{t} \), but their relative positions are reserved. So are their relative orientations, which can be explained in the polar coordinates. It is crucial to dynamic imaging.

\[
M_1(r - t) = F^{-1}\{\hat{I}_1(u)e^{-i\phi_1(u)}e^{i2\pi u \cdot r}\} \\
M_2(r - t) = F^{-1}\{\hat{I}_2(u)e^{-i\phi_2(u)}e^{i2\pi u \cdot r}\}.
\]

As shown in figure 1, when the object is moving, we captured a sequence of frames. We selected three frames and used MORE to recover the OTF phase, which was then used to directly deconvolve the images of all the frames with equation (5). The recovered images were simply stacked together to create a video without any image alignment process. Since the relative positions and orientations of the images are correctly recovered, the video displays the correct motion of the object. By adjusting the exposure time of the camera, we obtained three videos (400 frames each) at 25 Hz, 100 Hz and 200 Hz (see supplementary Movie1.avi, Movie2.avi and Movie3.avi (https://stacks.iop.org/NJP/22/093046/mmedia)). Figure 3 shows several frames of the video at 25 Hz. Please see appendix F for the details of the experiments and the other video samples.

Note that, in the process of each videos, MORE converged to the faithful OTF phase within 5 loops (equivalent to 15 separate iterations), taking less than 200 ms (with a Matlab code running on a Nvidia Tesla P100 GPU). MORE’s convergence speed is 10- to 1000-fold faster (in terms of iterations) than that of the HIO-ER (see appendices D and E for the comparison).
3.3. Imaging under low SNRs

For realistic imaging through turbid media, the detectable light is usually weak, resulting in a low SNR in the measurements. With an exposure time of 500 ms, the HIO-ER can recover reasonable images (see figure 9(c)). When the exposure time is down to 40 ms, the reconstruction with the HIO-ER becomes degraded. When the exposure time is 5 ms, the contrasts of the measured speckle patterns decrease to only $\sim 7\%$, and the HIO-ER fails to provide a reasonable results even after 120 000 iterations as shown in figure 4(c) (see appendix E for more details). In contrast, MORE still yields unique stable solutions (figure 4(d)). Moreover, as shown in figure 5, it converges from the first big iterative loop, and stably reaches its global minimum after the fifth big loop (less than 200 ms computation time in total). MORE exhibits a high stability in reconstruction and an ability of non-invasive imaging under a low SNR, making an big step close to a realistic application.

4. Discussion

Limitations of the speckle-correlation-based imaging method have been discussed in previous work [24]. One restriction is the practical limit on the complexity, $N$, of an object that can be recovered in one frame. $N$ is defined by the ratio of the object’s bright area to the area of the resolution cell [24]. Nevertheless, this problem can be ameliorated by reducing either the size of the localized illumination, or the area of the turbid media where light is collected from.

Another restriction is that, in order to obtain high-fidelity images, the illuminated region of the sample each time must be within an isoplanatic patch (memory effect range), as in adaptive optics, not just laterally but also axially. When the camera sensor is large enough, the lateral and axial speckle correlation ranges are
The convergence curves of the reconstructions for '2' to '6' in MORE. Data are taken with a 5 ms exposure time.

\[ z\lambda/\pi n L_{\text{eff}} \] [22] and \((2\lambda/\pi n)(z/D)^2 \) [26], respectively, where \( L_{\text{eff}} \) is the effective thickness of the medium. Any illumination originating from outside of this range contributes to noise in the measurement and deteriorate the image quality. \( L_{\text{eff}} = L - l_t \) if \( L \) is larger than the transport mean free path \( l_t \). However, in cases where \( L < l_t \), \( L_{\text{eff}} \) will be orders of magnitude smaller than the physical thickness \( L \), such as in biological tissues [24] and atmospheric turbulence. On the other hand, if the camera sensor is not large enough, the angular memory effect range \( \theta_{\text{ME}} \) will be limited by the size of the camera sensor \( D_{cs} \), \( \theta_{\text{ME}} = D_{cs}/d \).

Since MORE recovers images from the same OTF, it requires the turbid media between the object and the camera stays static during the measurement of one OTF. Thus, the exposure time for each frame must be much shorter than the varying timescale of the scattering media. For example, for a turbid media varying in a timescale of \( \sim 100 \text{ ms} \), we can record all the frames with an exposure time of 5 ms, which are then divided into groups to guarantee that all the frames in each group share the same OTF. Two adjacent groups should share a frame, which can be used to trace the movement of the object.

The operation of MORE has some similarities to ptychography [43], Fourier ptychography [44, 45] and super-resolution fluorescence imaging approaches, including speckle correlation imaging with a high-index scattering medium [46], as they all exploit joint information among multiple frames to optimize the iterative phase retrieval process. Our method, however, is different from the others in that it does not require overlap in the spatial (as described in appendix F) or in the Fourier domain, nor the low-resolution image as an intensity constraint. In an iterative process, MORE only employs two to five frames within an isoplanatic patch, whereas all frames are involved in an iterative computation for the other methods. Moreover, our method need no prior information on the illumination shape as in conventional structured illumination microscopy [47]. If combined with the scanning procedure proposed in high-resolution fluorescence imaging with a turbid medium [46], MORE may help to substantially improve the efficiency and fidelity of the image-recovery process.

In this work, we prove the OTF phase information is sufficient for imaging through optically-thick materials, and that it can be reconstructed in a phase retrieval process. It makes recovering images under low SNR situations possible. Meanwhile, its reliability and rapid convergence paves the way for real-time dynamic imaging through turbid media. We believe this idea will inspire many other applications with scattering materials such as super-resolution fluorescence imaging [46].

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 11503020).

Appendix A. The MTF

The imaging system’s PSF mainly depends on the physical property of the turbid media. It carries diffraction-limit information of the system with the speckle size being determined by the illuminated area.
Figure 6. Magnitude of the OTF, $|\tilde{S}(u, v)|$, for our imaging systems assuming an uniform illumination on the 6 mm circular diffuser. The curves are calculated by the van Cittert–Zernike theorem and plotted at the $v = 0$ plane for $C$ equal to 0.3, 0.6 and 1, respectively.

on the scattering medium. The theoretical upper cutoff frequency of the magnitude of the OTF is $f_{\text{upper}} \sim \frac{\mu m}{\lambda}$, where $n$ denote the refractive index behind the turbid medium. Note that, in reality, the system’s resolution can also be restricted by other effects, such as the size of the camera’s pixel. Moreover, it is worth mentioning that the system’s lower cutoff spatial frequency, $f_{\text{lower}}$, is affected by multiple factors including the size of the scatterers, the finite area of the detector’s sensor, and the uneven illumination background on the camera. In our experiment the last factor, the illumination background on the camera, determines $f_{\text{lower}}$.

With enormous numbers of random scatterers in the turbid medium, the PSF of the system at any one point is the result of a great many independent scattered wave vectors, and thus a Gaussian random variable [36, 37]. In ideal cases where more than thousands of speckles are captured by the camera’s sensor, for an arbitrary intensity distribution of light transmitted through the scattering layer $T(\xi, \eta)$, the autocorrelation of the PSF is predicted by the van Cittert–Zernike theorem [36]:

$$S \ast S(x, y) = \bar{S}^2 \left[ C^2 |\gamma(x, y)|^2 + 1 \right]$$  \hspace{1cm} (A1)

where

$$\gamma(x, y) = \frac{e^{-i\phi} \int \int T(\xi, \eta) \exp \left[ i \frac{2\pi}{\lambda} (x\xi + y\eta) \right] d\xi \, d\eta}{\int \int T(\xi, \eta) d\xi \, d\eta}$$  \hspace{1cm} (A2)

$\bar{S}$ is the average intensity of the PSF. $\xi$ and $\eta$ represent the coordinates on the diffuser plane. $\phi$ is the phase factor [36]. $C$ is the contrast of the PSF, which is less than 1. Its value is influenced by various factors including the noise level in the measurement, the polarization and the bandwidth of the light source. Based on equations (A1) and (A2), if the transmitted light through the medium has an uniform circular cross section, the autocorrelation of the PSF will be a Gaussian-like function (an airy disk) with a background [26, 48]. In Fourier domain,

$$|\tilde{S}(u, v)|^2 = \bar{S}^2 \left[ C^2 |T \ast T|(u, v) + \delta_0(u, v) \right],$$  \hspace{1cm} (A3)

Thus $|\tilde{S}(u, v)|^2$, as the Fourier transform of the autocorrelation of the PSF, is also a Gaussian-like function except for a sharp peak in the very middle. In figure 6, curves of $|\tilde{S}(u, v)|/\bar{S}$ are plotted at the $v = 0$ plane assuming an even intensity distribution on the circular-shape diffuser with different values of $C$. Note that, $|\tilde{S}(u, 0)|$ has nonzero values only when $|u|$ is below $\sim f_{\text{upper}}$. In our experiments, we have a contrast of the PSF $C \lesssim 30\%$.

Appendix B. The determination of a support

The region of the object can be estimated from the power spectrum of a captured pattern. The size of a mode of the spatial frequency is inversely proportional to the region of the object. In our experiment, the size of a mode is around $17 \times 29$ pixels in the discrete Fourier domain. The region of the object $\Gamma$ is set as a
Figure 7. The recovered images with different sizes of $\Gamma$. 0% represents the size of $128 \times 70$ pixels, estimated from power spectrum of the captured pattern. The others are of the percentage changes with respect to the estimation.

Figure 8. The convergency of the reconstruction after 10 big loops VS different size of $\Gamma$. rectangle of $(2048/17) \times (2048/29) \approx 128 \times 70$ pixels, where $2048 \times 2048$ pixels is the dimension of the captured pattern. We also investigate how the accuracy of the estimated region affects the reconstruction. Figure 7 shows the reconstruction results VS different estimated region. Figure 8 shows the relevant convergencies. These results indicate that, 1) the estimation from the power spectrum is faithful; 2) the reconstruction gradually degrades when $\Gamma$ is getting smaller than its true size. 3) MORE is less sensitive to a bigger $\Gamma$. When $\Gamma$ is set to be twice bigger than the object’s true size, MORE still provides a reasonable result.

Appendix C. Comparison of reconstructions for the data with a 500 ms exposure time

The speckle patterns were captured with a 500 ms exposure time. In order to compare the relative positions and orientations of the recovered images, we recorded each position of the object while measuring it. After all the measurements, we restored the location of each object and took its direct image. Under the 500 ms exposure time, the HIO-ER can recover good images. Nevertheless, we still need to run more than 10,000 iterations for each object, and choose the best result from them. Figure 9(c) shows that, when using the
HIO-ER to individually recover an object from a single frame, the relative positions and orientations of different objects are lost. In contrast, MORE preserves their relative positions and orientations, as shown in figure 9(b).

Appendix D. Reconstruction for the data taken at 25 Hz (40 ms exposure time)

We here compare the performance of MORE with that of the HIO-ER, which uses a single speckle pattern to recover an object [23, 24]. The speckle patterns were captured at 25 Hz (40 ms exposure time) while the object is moving. Figure 10 shows the convergence curves for object ‘2’ and ‘5’ in the first 200 iterations of the HIO. The normalized difference between a recovered image $O_{\text{rec}}(x, y)$ and the direct image $O_{\text{dir}}(x, y)$ is calculated with equation (D1):

$$\frac{\sum_{x,y} |O_{\text{rec}}(x,y) - O_{\text{dir}}(x,y)|}{\sum_{x,y} |O_{\text{dir}}(x,y)|}. \quad (D1)$$

Note that, at $K = 0$, all the recovered images are calculated with an initial random guess of the OTF. In each calculation of the difference, each recovered image was rotated to the same orientation as that of the direct image, and shifted to the best position to have a maximal overlap with the direct imaging. This process guarantees the validity of the comparison.

The convergence curves indicate that, the reconstructions oscillate around local minima. The convergence curves of ‘3’, ‘4’ and ‘6’ also exhibit a similar oscillation. Therefore, it is necessary to choose the best result for an object in all the iterations by looking for the closest Fourier magnitude to $|\tilde{I}_{\text{pro}}(u,v)|$, i.e., the minimum of...
Figure 11. The recovered images of ’2’ to ’6’ at the several iterations. ’Best’ show the best result after 50 000 iterations for each object. Data are taken with a 40 ms exposure time.

Figure 12. The convergence curves of the reconstructions for ’2’ to ’6’. (a) The first 10 iterations; (b) the first 200 iterations. Data are taken with a 40 ms exposure time.

$$\sum_{u,v} \left| \hat{O}_{rec}(u,v) \right| - \left| \hat{I}_{pro}(u,v) \right| \sum_{u,v} \left| \hat{I}_{pro}(u,v) \right|, \quad (D2)$$

where $\hat{O}_{rec}(u,v)$ is the Fourier magnitude of the recovered image. This chosen criterion is not very reliable. Most of the time, we manually selected the best ones based on the prior information (their direct images). Figure 11 exhibits the recovered image at the iterations of $j = 0, 5, 10, 15, 40, 45$, as well as the best result chosen from 50 000 iterations for each object.

We perform MORE with the same data set. The initial condition is also the same as that used in the HIO-ER. Figures 12 and 13 shows convergence speeds and all the recovered images after each big loop, respectively. MORE converges to the right results since the first big loop. The convergence reaches its global minimum after the third big loop.
Appendix E. Reconstruction for the data taken at 200 Hz (5 ms exposure time)

When the object was moving, we captured the speckle patterns at 200 Hz with a 5 ms exposure time. The data are very noisy. The average contrast of the speckle patterns is $\sim 7\%$.

The ‘normalized difference’ represents the differences between the recovered images and their direct images, calculated in the same way as in the previous section. In figure 14, the convergence curves show similar oscillations. To get a faithful reconstruction, we ran more than 120,000 iterations for each object: 500 trials ran with 25 different beta factors and 20 different initial random conditions; each trial included 200 iterations of the HIO and 50 iterations of the ER. Then the best reconstruction was chosen by looking for the closet Fourier magnitude to the $|\tilde{I}^{\text{pro}}(u,v)|$. As shown in figure 15, even after 120,000 runs, none of the results exhibits a sign of convergence.

For comparison, we show the performance of MORE for the same data. We also used the same initial guesses. As shown in figures 16 and 17, MORE converges to the right results since the first big loop. The convergence reaches its global minimum after the fifth big loop.

Appendix F. Experimental details of dynamic imaging

The experimental setup for dynamic imaging is shown in figure 1 of the main text. We tested two objects.
Figure 15. The recovered images of ‘2’ to ‘6’ at the several iterations of the HIO for reconstruction with a single frame. The beta factor is 0.75. ‘Best’ shows the best result chosen from 120 000 iterations for each object. Data are taken with a 5 ms exposure time.

Figure 16. The convergence curves of the reconstructions for ‘2’ to ‘6’ in MORE. Data are taken with a 5 ms exposure time.

One (dynamic object 1) had a bigger size (each numeral is of ∼5 mm), which was easy to be bended and attached on the rotatable cylinder. Another (dynamic object 2) has a much smaller size. Each numeral is of 0.5 × 0.4 mm. So, we were able to investigate the resolution down to μm level. But it was difficult for us to make a cylinder with such a small size. We therefore used a moving plate (containing numerals) and a screen to mimic the behavior of the rotating cylinder: one numeral can be seen each time in the field of view of the measurement, as shown in figure 18.

For dynamic object 1, we set the distance from the object to the diffuser as $z \approx 2$ m, the distance between the diffuser and the camera as $d \approx 200$ mm. A sequence of frames was captured by the camera while the cylinder was rotating. We selected three frames corresponding to three different numerals to
Figure 17. All the recovered images of the first 4 big loops in MORE for the data taken at 200 Hz (5 ms exposure time).

Figure 18. (a) Dynamic object 1. (b) Dynamic object 2

Figure 19. The recovered images of the numerals on dynamic object 1.

recover the OTF phase. All the images were then directly computed with this recovered OTF phase using equation (6) of the main text. Figure 19 shows the results.

For dynamic object 2, we reset the experimental parameters to $z \approx 200$ mm and $d \approx 300$ mm. We performed the dynamic imaging in the same way as for dynamic object 1. After we acquired all the images at different moments of the object, we simply stacked all the images together and generate a video without any image alignment process: each frame is just a recovered image. Since the preservation of the relative positions and orientations, the video displays the real motion of the object. By setting the exposure time of the camera to 40 ms, 10 ms and 5 ms, we obtained the video imaging at the frame rate of 25 Hz, 100 Hz and 200 Hz, respectively. Figure 20 shows some frames of the videos. Please supplementary Movie1.avi, Movie2.avi and Movie3.avi for the whole videos.

Note that, the dimension of each frame was $2048 \times 2048$. The computation time for the OTF phase was less than $\sim 200$ ms with a three years old GPU (Nvidia Tesla P100 GPU). The reconstruction for each frame contains two Fourier transform calculations, which cost several ms in our laptop.
Figure 20. Samples of video frames reconstructed by MORE. (a) Samples of the 25 Hz video for imaging through turbid media. See supplementary Movie1.avi. (b) Similar to (a) for the 100 Hz video. See supplementary Movie2.avi. (c) Similar to (a) for the 200 Hz video. See supplementary Movie3.avi. The OTF phases are derived for each video individually.

ORCID iDs

Hui Chen https://orcid.org/0000-0002-2735-1068

References

[1] Gibson A P, Hebden J C and Arridge S R 2005 Phys. Med. Biol. 50 R1
[2] Koenderink A F, Lagendijk A and Vos W L 2005 Phys. Rev. B 72 153102
[3] Hardy J W 1998 Adaptive Optics for Astronomical Telescopes vol 16 (Oxford: Oxford University Press)
[4] Tang J, Germain R N and Cui M 2012 Proc. Natl Acad. Sci. 109 8434
[5] Rueckel M, Mack-Bucher J A and Denk W 2006 Proc. Natl Acad. Sci. 103 17137
[6] Hee M R, Swanson E A, Izatt J A, Jacobson J M and Fujimoto J G 1993 Opt. Lett. 18 950
[7] Niedre M J, de Kleine R H, Aikawa E, Kirsch D G, Weissleder R and Ntziachristos V 2008 Proc. Natl Acad. Sci. 105 19126
[8] Huang D et al 1991 Science 254 1178
[9] Kang S et al 2015 Nat. Photon. 9 253
[10] Xu X, Liu H and Wang L V 2011 Nat. Photon. 5 154
[11] Wang L V and Hu S 2012 Science 335 1458
[12] Liu Y, Lai P, Ma C, Xu X, Grabar A A and Wang L V 2015 Nat. Commun. 6 5904
[13] Popoff S, Lerosey G, Fink M, Boccara A C and Gigan S 2010 Nat. Commun. 1 626
[14] Edrei E and Scarcelli G 2016 Optica 3 71
[15] Vellekoop I M and Mosk A P 2007 Opt. Lett. 32 2309
[16] Vellekoop I M, Lagendijk A and Mosk A P 2010 Nat. Photon. 4 320
[17] Katz O, Small E and Silberberg Y 2012 Nat. Photon. 6 549
[18] Vellekoop I M and Mosk A P 2008 Phys. Rev. Lett. 101 120601
[19] Hsieh C-L, Pu Y, Grange R, Laporte G and Psaltis D 2010 Opt. Express 18 20723
[20] Wang D, Zhou E H, Brake J, Ruan H, Jang M and Yang C 2015 Optica 2 728
[21] Feng S, Kane C, Lee P A and Stone A D 1988 Phys. Rev. Lett. 61 834
[22] Freund I, Rosenbluh M and Feng S 1988 Phys. Rev. Lett. 61 2323
[23] Bertolotti J, van Putten E G, Blum C, Lagendijk A, Yos W L and Mosk A P 2012 Nature 491 232
[24] Katz O, Heidmann P, Fink M and Gigan S 2014 Nat. Photon. 8 784
[25] Labeyrie A 1970 Astron. Astrophys. 6 85
[26] Freund I 1990 Phys. A 168 49
[27] Yuan Y and Chen H 2017 Non-invasive dynamic or wide-field imaging through opaque layers and around corners (arXiv:1712.08576 [physics.optics])
[28] Edrei E and Scarcelli G 2016 Sci. Rep. 6 33558
[29] Xu X, Xie X, He H, Zhuang H, Zhou J, Thendiyammal A and Mosk A P 2017 Opt. Express 25 32829
[30] Li L, Li Q, Sun S, Lin H-Z, Liu W-T and Chen P-X 2018 Opt. Lett. 43 1670
[31] Wang Z, Jin X and Dai Q 2018 Sci. Rep. 8 9088
[32] Bucchi A, Donatelli M and Ferri F 2018 J. Sci. Comput. 74 786
[33] Lu D, Liao M, He W, Pedrini G, Osten W and Peng X 2020 Opt. Lasers Eng. 124 105815
[34] Mukherjee S, Vijayakumar A, Kumar M and Rosen J 2018 Sci. Rep. 8 1134
[35] Liao M, Lu D, Pedrini G, Osten W, Situ G, He W and Peng X 2019 Sci. Rep. 9 7165
[36] Goodman J W 2000 Statistical Optics (Wiley Classics Library) (New York: Wiley) p 572
[37] Goodman J W 1975 Laser Speckle and Related Phenomena (Berlin: Springer) pp 9–75
[38] Korff D 1973 J. Opt. Soc. Am. 63 971
[39] Richardson W H 1972 J. Non-Soc. Am. 62 55
[40] Lucy L B 1974 Astron. J. 79 745
[41] Fienup J R 1982 Appl. Opt. 21 2758
[42] Ancora D, Di Battista D, Giasafaki G, Psycharakis S E, Liapis E, Ripoll J and Zacharakis G 2018 Methods 136 81
[43] McCullum B C and Rodenburg J M 1993 J. Opt. Soc. Am. A 10 231
[44] Zheng G, Horstmeyer R and Yang C 2013 Nat. Photon. 7 739
[45] Dong S, Nanda P S, Shiralikar R, Guo K and Zheng G 2014 Opt. Express 22 20856
[46] Yilmaz H, van Putten E G, Bertolotti J, Lagendijk A, Yos W L and Mosk A P 2015 Optica 2 424
[47] Gustafsson M G L 2000 J. Microsc. 198 82
[48] Freund I, Rosenbluh M and Berkovits R 1989 Phys. Rev. B 39 12403