Feedback-based wavefront shaping

Ivo M. Vellekoop

Biomedical Photonic Imaging group, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

Abstract: Light scattering was thought to be the fundamental limitation for the depth at which optical imaging methods can retain their resolution and sensitivity. However, it was shown that light can be focused inside even the most strongly scattering objects by spatially shaping the wavefront of the incident light. This review summarizes recently developed feedback-based approaches for focusing light inside and through scattering objects.

©2015 Optical Society of America

OCIS codes: (290.0290) Scattering; (030.6140) Speckle; (030.6600) Statistical optics.

References and links


1. Introduction

Telescopes, cameras, and microscopes all rely on the notion that light propagates along a straight path. However, when light propagates through a turbid medium, its direction gets scattered by scattering on microscopic particles and other inhomogeneities. As a result, no sharp focus is formed. For long, it was believed that light scattering posed a fundamental limitation on the penetration depth and resolution of all optical methods [1]. The paradigm to use ballistic (non-scattered) light for imaging has been pushed to extreme limits with the ability to separate ballistic and scattered light through spatial selection [2–4] or coherence gating [5], and through the advances of multi-photon microscopy [6]. Even so, these methods still rely on ballistic light to form an image. Since the amount of ballistic light decreases exponentially with depth, increasing the penetration depth of a microscope without compromising its resolution is exponentially hard. Even with the most advanced multi-photon microscope, the typical penetration depth in biological tissue is less than a millimeter [7]. For long, the only options to image at larger depths were diffuse optical tomography and related methods; approaches that unfortunately suffer from a severely reduced resolution [8, 9].

In the last few years, however, it is becoming increasingly clear that scattering needs not be a fundamental limitation for imaging. In 2007 Allard Mosk and I developed a technique,
now called wavefront shaping, that can be used to focus light through [10] or even inside scattering objects [11] (see Fig. 1). Our message was that light scattering is not a fundamental limitation. Even though light scatters in a complex and initially unknown way, the scattering process is linear and entirely deterministic as long as the medium is static. By shaping the incident wavefront in just the correct way, light can be focused anywhere, even deep inside the most strongly scattering materials.

Fig. 1. Principle of wavefront shaping. a) when a scattering sample is illuminated with a plane wave, the scattered light forms a disordered interference pattern known as laser speckle. b) measured intensity of speckle pattern (logarithmic color scale). c) a spatial light modulator is used to shape the wavefront of the incident beam; the light modulator is programmed using feedback from a detector placed behind the object. After completion, the feedback algorithm has found the optimum solution for focusing light through the sample the intensity in the target has increased by over three orders of magnitude (d). e) by combining feedback from multiple points, simple images can be projected through the object. f) phase of the incident wavefront that was used for measuring (e), no correlations can be observed in the incident field, indicating that the sample fully scrambles the incident wave. Image adapted from Vellekoop & Mosk [10].

The ability to focus light at any desired position has a huge potential for applications. The field of wavefront shaping is under rapid development, and proof-of-concept applications in deep tissue microscopy [12–15], endoscopy [16–19], optical trapping [20], super-resolution imaging [21, 22], nano-positioning [23], and cryptography [24–26] have been demonstrated.

Even though light can be focused anywhere once the correct incident wavefront is known, finding this matching wavefront is far from trivial. When it is allowed to place a detector in the desired focus, a simple feedback scheme (Fig. 1) can be used. For most applications, however, a different approach is required. In the last couple of years, a large variety of approaches and algorithms for focusing light in the presence of scattering have been demonstrated. In this review, I focus only on the category of feedback-based optical wavefront shaping. Therefore, I will not attempt to cover the rapidly expanding field of optical phase conjugation, nor the foundational work utilizing other wave modalities, such as ultrasound and radio waves. Recent reviews of these fields can be found in Horstmeyer et al. [27] and Mosk et al. [28], respectively.

First, I will introduce the concept of feedback-based wavefront shaping and some of the fundamental properties of this technique. Then, I will review the different algorithms that can be used for finding the wavefront, and discuss different options for obtaining a feedback signal in the first place. After that, I touch on some of the wave correlations that play a role in wavefront shaping, and briefly link to related research fields. This review is concluded with an outlook of future applications.
2. Feedback-based wavefront shaping

When a turbid medium is illuminated with coherent light, the transmitted light forms a disordered interference pattern that is known as laser speckle (see Fig. 1(b)). For a given turbid sample it is not possible to predict the speckle pattern in advance. Therefore, most work on multiple scattering is aimed at finding a statistical description for an ensemble of similar samples (e.g. [29, 30]). Often, the term ‘random scattering’ is used. This term, however, is somewhat misleading since for a given sample light scattering is completely deterministic.

Wavefront shaping exploits the fact that light scattering is linear and fully deterministic. Light transport through any linear medium – including absorbing media and magneto-optic materials where time-reversal symmetry is broken – can be described using a general matrix approach:

\[
E_b = \sum_a t_{ba} E_a
\]

where \(a\) are the indices of the components, or modes, of the incident field \(E_a\) in some arbitrary basis, and, similarly, \(b\) are the indices of the components of the transmitted field \(E_b\), potentially in a different basis. The summation is over all incident modes. For simplicity, we assume that both bases are orthonormal, although this is not strictly required. The elements \(t_{ba}\) of the transmission matrix (TM) describe scattering in the sample. The matrix approach is a convenient method that is used extensively in multiple scattering theory [31], where the basis vectors of the incident and outgoing field are called ‘channels’. Any orthogonal basis can be used and, when desired, the basis can be chosen to include both polarization states of the light. The ‘transmission matrix’ Eq. (1) can be replaced by a different matrix without loss of generality to model e.g. reflection, or transport to a point inside the sample, or – in fact – any other linear process.

In its simplest form, feedback based wavefront shaping is concerned with solving the following optimization problem: what is the incident field \(E_a\) that maximizes transmission into a desired output mode \(\beta\). Here, mode \(\beta\) can represent e.g. a focus, a plane wave, or any other desired field pattern.

Fortunately, this optimization problem is very easy to solve for a linear system when the TM is known. Defining the ‘intensity’ in mode \(\beta\) as \(I_\beta = |E_\beta|^2\), we have from Eq. (1)

\[
I_\beta = \left| \sum_a t_{ba} E_a \right|^2 \leq \sum_a |t_{ba}|^2 \sum_a |E_a|^2
\]

where the Cauchy-Schwartz inequality was used in the second step. The equality only holds when \(E_a \propto t_{ba}^*\). Assuming that we are not allowed to change the total incident power, \(I_{in} = \sum_a |E_a|^2\) remains constant. Therefore, the incident field that maximizes transmission into mode \(E_\beta\) is simply given by

\[
E_a = E_0 t_{ba}^*
\]

with \(E_0\) a normalization constant to ensure that the total incident power remains fixed. This solution represents a global maximum that is unique up to an arbitrary phase factor. Interestingly, this solution is exactly the phase conjugate of a wave propagating from \(\beta\) back to \(a\) (see Section 6.2 for a brief discussion on the relation with phase conjugation).
It should be noted that it is common practice to use the term ‘intensity’ liberally. In a matrix representation there are no spatial coordinates and ‘intensity’ is a quantity that is proportional to the total power (Watts) in the mode: ‘intensity’ and ‘power’ are equivalent. In a continuous field description, however, the distinction must be made, and I will use the terms ‘intensity’ and ‘power’ consistently with a continuous field representation (see e.g. [32],), where power (W) is intensity (W/m²) integrated over area.

2.1. Enhancement

One of the most common figures of merit is the enhancement $\eta$. It is defined as

$$\eta = \frac{\hat{I}_\beta}{\langle I_\beta \rangle}$$

where $\hat{I}_\beta$ is the intensity in the focus after optimizing the incident wavefront, and $\langle I_\beta \rangle$ is the reference intensity. Strictly speaking, the correct way to determine the reference intensity is as the intensity in $\beta$ measured for the same incident wavefront that was used for measuring $\hat{I}_\beta$, but averaged over the ensemble of possible samples. Ensemble averaging can usually be replaced by translation of the sample, or by time averaging (for dynamic media). Also, one may use the average background intensity around the optimized focus as a reference, in which case $\eta$ equals the signal to background ratio (SBR), or simply contrast.

In experiments, the total power that is incident on the sample surface often depends strongly on the pattern that is displayed on the light modulator, mainly because of diffraction and vignetting in the light path. Therefore, it is not correct to use the initial (non-optimized) intensity in $\beta$ as a reference. The enhancement and contrast are not sensitive to this type of systematic error, making them proper figures of merit. Note, however, that when a large fraction of the incident modes is controlled with a high accuracy, wavefront shaping increases the background intensity as well as the intensity in the focus [33], making the SBR contrast a less suitable figure of merit in this special case.

In a strongly scattering sample, the elements of a row of the TM are statistically independent and follow a circular Gaussian distribution [30]. Under these circumstances, the maximum enhancement that can be achieved is [33]

$$\eta = \alpha (N-1) + 1$$

where $N$ is the number of independently controlled input channels. The factor $\alpha \in [0,1]$ depends on the type of light modulation that is used, and on the intensity distribution over the input channels. For an ideal light modulator that modulates both phase and amplitude of the light, $\alpha = 1$. For other cases, $\alpha$ can be calculated by taking the average inner product of the ideal field and the actual field [32, 33]. A few common scenarios are listed in Table 1. In practice, inhomogeneous illumination of the SLM, cross talk between elements of the SLM, and measurement noise will further reduce the enhancement [34, 35]. Typical enhancements reached in experiments currently range from $10^2$ to $>10^3$, depending mainly on the number of control elements and the stability of the sample (see Section 5.1).
Table 1. Relative maximum enhancement $\alpha$ for different wavefront shaping methods.

<table>
<thead>
<tr>
<th>Light modulator type</th>
<th>$\alpha$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude and phase</td>
<td>1</td>
<td>[33]</td>
</tr>
<tr>
<td>Phase only</td>
<td>$\pi/4$</td>
<td>[10]</td>
</tr>
<tr>
<td>Intensity only</td>
<td>$1/(2\pi)$</td>
<td>[34]</td>
</tr>
<tr>
<td>Binary phase</td>
<td>$1/\pi$</td>
<td>[36, 37]</td>
</tr>
</tbody>
</table>

2.2. Total transmission into the focus

Even though wavefront shaping techniques can increase the intensity at a given point, it is theoretically impossible to focus all incident light to that point. The fraction of the incident power that reaches the focus after optimization is usually far less than one percent. From Eq. (2) we can directly see:

$$\frac{I_\beta}{I_{in}} \leq T_\beta \quad \text{with} \quad T_\beta = \sum_n |f_{\beta n}|^2$$

(6)

Here, $T_\beta$ is equal to the total (power) transmission into mode $\beta$, averaged over all possible incident wavefronts. For phase conjugation experiments this relation represents a simple symmetry, in this context $T_\beta$ can be interpreted as the fraction of the source power that reaches the phase conjugation system. If, for example, a phase conjugation system collects 0.1% of the light from the source, no more than 0.1% of the phase conjugated light can be sent back to the reconstructed focus. The rest of the light will partially be reflected, and partially forms a background speckle around the focus. Usually $T_\beta \ll 1\%$ due to a combination of factors: reflected light is not collected, usually only a single polarization is used, the system has a finite numerical aperture and a finite extent, the sample may absorb some of the light, etc. For feedback-based wavefront shaping the situation is even worse, because the number of control element is usually limited to a few thousand for performance reasons. Fortunately, for most applications it is not essential to put a large fraction of the total energy in the focus: achieving a good contrast is sufficient.

2.3. Overall transmission

An intriguing aspect of wave scattering is the existence of open transport channels. In 1984, Dorokhov predicted that for each non-absorbing scattering sample it is possible to construct a wave that is transmitted completely (i.e., experiences zero reflection) [38]. The existence of such open channels is responsible for a large variety of mesoscopic phenomena, including universal conductance fluctuations [31]. Originally studied mostly in the context of electron scattering, the phenomenon universally applies to wave scattering. Important pioneering work was done microwaves and ultrasound [28], for which the field of the wave can be detected over a large bandwidth relatively easily. For light, an interferometric technique is needed to measure the phase of the wave, and experiments are usually conducted with monochromatic waves. However, the huge advantage of using light is the availability of spatial light modulators (SLMs) with millions of degrees of freedom and detectors with millions of pixels.

The wavefronts corresponding to open channels are singular vectors of the TM of a closed system. In this context, ‘closed’ means that all degrees of freedom of the incident field are controlled, and all degrees of freedom of the transmitted field are resolved, without any light being absorbed or leaking from the system. In an optical experiment, only a subset of the TM of the sample can be determined; usually limited by the numerical aperture of the microscope.
objectives that are used. Still, the total energy transmission through a sample can be increased if a large fraction of the degrees of freedom in the incident field is controlled \[33, 39, 40\]. Also, if only a small subset of the TM is known, it is also possible to increase the total transmitted power locally at the detector surface \[41\].

3. Algorithms

The TM for a given object is not known a-priori: it depends on the microscopic refractive index distribution in the sample, which will be different for each specimen. In order to focus light to a point, there are two types of approaches: 1) first measure the TM elements and then calculate the ideal field using Eq. (3), or 2) find the optimum field using an iterative optimization algorithm. This section contains a description of the most commonly used approaches for forward wavefront shaping. All algorithms described here are aimed at maximizing $\eta$, and they are designed for the situation where the light is completely diffuse.

![Fig. 2. Stepwise sequential algorithm. This algorithm measures a single row of the TM of a scattering sample. The light modulator is divided into a series of segments, and the phase of each consecutive segment is swept from 0 to $2\pi$ (depicted by the crossed out segment) while monitoring the feedback signal. This way, the phase and amplitude of the elements of the TM row are measured. After all measurements are completed, the wavefront is updated (far right) to display the optimum wavefront for focusing. Image reproduced from \[42\].](image)

3.1. Transmission matrix measurements

This class of algorithms measures the elements of the TM first, and then calculates the optimal wavefront for focusing light onto a point or a set of points. Experimentally, the simplest approach to measure the matrix elements is to use the sample itself as an interferometer. Part of the incident field is kept static and serves as a co-propagating reference beam; resulting in a static speckle pattern (the reference field) at the back of the sample. The phase of the matrix elements is determined relative to the phase of this reference field.

This approach is used in the ‘stepwise sequential’ algorithm (see Fig. 2) For this algorithm a phase-only spatial light modulator is subdivided into $N$ segments, and the phase $\phi_\alpha$ of each segment $\alpha$ is cycled between 0 and $2\pi$ consecutively. At the detector, the light from the modulated segment $E_{\beta\alpha}$ interferes with the light originating from all other segments $E_{\beta\beta}$.

$$I_\beta(\phi_\alpha) = |E_{\beta\alpha} + E_{\beta\beta}|^2 = |E_{\beta\alpha}|^2 + |E_{\beta\beta}|^2 + 2|E_{\beta\alpha}E_{\beta\beta}| \cos(\phi_{\beta\alpha} + \phi_\alpha)$$

With $\phi_{\beta\alpha}$ the phase of the transmission matrix element $t_{\beta\alpha}$ relative to the unknown phase of the reference field. The relative values of $t_{\beta\alpha}$ can be extracted from this measurement by fitting a cosine. By repeating the measurement for each segment, all elements of a row of the TM can be obtained up to a single common pre-factor. After determining the complete row of matrix elements, the light modulator is programmed such that $\phi_{\beta\alpha} = -\phi_{\beta\alpha}$, which causes constructive interference (focusing) at point $\beta$.

Using the field from the non-modulated segments as a reference beam is only allowed when a small fraction of the segments is modulated at a time, so that background field $E_{\beta0}$ will be nearly equal for each segment. Unfortunately, in this case the signal to noise ratio (SNR) of the measurement may not be optimal. To optimize the performance in the presence
of shot noise and technical noise, a pre-optimization with a low $N$ can be used first to increase the reference intensity [35]. A different approach to improve the SNR of the measurements is to modulate multiple segments at once, for example by choosing a Hadamard basis. In this case, a different approach should be used, such as reserving a part of the light modulator for generating the co-propagating reference beam [43, 44]. Also, instead of measuring the elements of the TM sequentially, it is possible to modulate each pixel with a different frequency and Fourier transform the detector signal to separate the different signals [45, 46].

When the point detector in Fig. 1 is replaced by a camera, it is possible to measure multiple rows of the TM at once [43]. This technique is extremely powerful as it allows for transmitting images [47, 48], direct imaging [22, 49], raster scanning fluorescence microscopy [16], and maximizing the transmission to a given large-area target [41]. Also, it allows a determination of the eigenvalues of the TM [39, 41, 43], the statistical distribution of which is of fundamental interest in mesoscopic scattering theory (e. g [31]). When it is sufficient to know the transmission matrix up to a common phase factor for each row, a co-propagating reference beam like described above can be used. If the relative phase between different rows is important, an external reference beam can be used instead (see e. g [41]).

![Fig. 3. Random partitioning algorithm. In this iterative approach, half of the segments of the light modulator is chosen at random (crossed squares). The phase of these selected segments is varied until the maximum intensity is reached. After that, the iteration is repeated with a different subset of selected segments. Image reproduced from [42].](image)

### 3.2. Iterative algorithms

Instead of measuring the matrix elements directly, a variety of iterative approaches can be used to find the optimal wavefront for creating a focus. Where TM approaches only update the wavefront once all measurements are complete, iterative methods update the wavefront after each single step. Iterative approaches have two advantages: 1) the intensity in the target starts increasing right away. 2) in many cases the SNR is better than for TM methods [50].

The fact that the target intensity starts increasing right away is especially important when the medium is dynamic. TM algorithms don’t perform well, if at all, if the medium changes before all measurements are completed. Therefore, the number of segments $N$ needs to be adjusted to the speckle decorrelation time of the medium. If the decorrelation time is not known, or not constant, it may be favorable to use an iterative approach to reach the highest possible enhancement without adjusting any parameters for the tissue dynamics [42].

A simple iterative approach is illustrated in Fig. 3. In this random partitioning algorithm, the pixels of the light modulator are randomly subdivided into two partitions of equal size, and the relative phase of the partitions is adjusted to maximize the feedback signal. Then the procedure is repeated for a different random subset, ad infinitum. When this algorithm is started, the intensity in the focus increases rapidly and the performance is good even with a low initial signal to noise ratio. However, after several steps the algorithm starts slowing down since the partitions are not truly orthogonal. Moreover, the required phase corrections shrink over time, so the phase should be determined with a very high accuracy [42]. Any error in determining the phase can be devastating: if a single noisy measurements gives an error of $\pi$, the focus disappears abruptly. Therefore, the algorithm should be able to ‘backtrack’ to a previously stored state if the focus is lost.
An interesting category is formed by genetic algorithms, first employed in this context by Conkey et al. [50]. Genetic algorithms start with a pool of randomly generated wavefronts and select the ones with the highest enhancement (see Fig. 4). These ‘winning’ wavefronts are mixed randomly and modified to generate the next generation of wavefronts. In simulations, genetic algorithms are very robust to additive noise (see Fig. 5). Like the random partitioning algorithm, convergence does slow down after a certain number of iterations.

3.3. Discussion

Regardless of what algorithm is used, ideally all methods converge to the unique global optimum given by Eq. (4). For TM methods, the global optimum is reached after performing a single measurement for each row of the TM. The iterative methods generally do not converge to the maximum in an entirely deterministic manner, and their convergence slows down significantly after several iterations. For a strongly scattering sample, all $N$ incident modes are statistically independent. Therefore, any given algorithms will require at least $O(N)$ measurements to achieve an enhancement of $N$. The actual number typically ranges from $N$ to $10N$. As such, iterative methods are not faster than TM methods; their main advantages are the lower sensitivity to noise [50] and their quick recovery after sample movement [42].
3.4. Extended targets

The algorithms described above are able to find the optimal wavefront for focusing to a spot that is the size of a single speckle. However, when the target area is larger than a single speckle, these methods are not guaranteed to find the optimum solution. The behavior of the stepwise sequential algorithm in this situation can be analyzed easily [42]: the algorithm is expected to find the wavefront that maximizes transmission into the effective target ‘mode’ \( E_{r}^{\text{eff}}(r) \equiv D(r)E_0(r) \), with \( D(r) \) the envelope of the detection sensitivity, and \( E_0(r) \) the original speckle field at the start of the algorithm. Therefore, the end result depends on the starting conditions of the algorithm. This behavior, however, is likely to depend on the algorithm that is used. For instance, it was observed that genetic algorithms caused the light to focus onto the brightest speckle [51].

Even though it is still possible to increase the amplitude in \( E_{r}^{\text{eff}} \) by the same factor as for a point target, the SBR contrast will be lower. For instance, consider the situation where the envelope \( D(r) \) spans an area that supports \( M \) optical modes. In this case, increasing the intensity of only one of these modes by a factor \( \eta \) will only cause an overall increase of the target signal of a factor of \( \eta/M \) [42].

The difficulty in focusing to multiple spots at once is a fundamental problem that plays a significant role in wavefront shaping with photoacoustic feedback or fluorescence feedback when the target signal comes from an area that spans many optical modes. The same problem is seen in time reversal of ultrasound encoded light, where typically \( N = 10^5 \) and \( M = 10^4 \), resulting in a contrast of around 20 [12, 13, 15].

4. Feedback

All algorithms discussed above rely on the availability of a feedback signal to optimize. In the simplest case we can use a camera or point detector placed behind the sample. However, if we want to use wavefront shaping for imaging inside turbid media, a feedback signal from inside the medium is needed. I will now review different approaches for obtaining such a signal.

4.1. Fluorescence

A small fluorescent particle can be used as a probe to locally measure the amount of excitation light. The total amount of fluorescent emission is a direct measure for the
excitation intensity at the location of the particle. By using the total fluorescence emission as a feedback signal, light was focused onto a fluorescent nanoparticle embedded at a depth of ~20 µm inside strongly scattering ZnO pigment, reaching an enhancement of > 20 [11]. Care must be taken that the labelled volume is small (see Section 3.4) and that the autofluorescence of the medium is sufficiently low to detect the signal from the probe. Since it is often possible to fluorescently label structures of interest, this method may be used for focusing light on labelled cells or organelles.

Tang et al. [52] used a two-photon fluorescence signal as a feedback for focusing light inside turbid tissue. In adaptive optics for two-photon microscopy, the two-photon fluorescence signal is often used as a feedback signal to optimize the image quality. However, Tang et al. demonstrated that this approach is even feasible when scattering is so bad that there is no image to start with.

4.2. Non-linear feedback

When a two-photon fluorescence signal is used, the feedback signal is proportional to the square of the intensity in the medium. This property adds two interesting and useful characteristics to the wavefront shaping process. First of all, the feedback signal responds to an improved spatial confinement of the signal, as well as to an improved temporal confinement. In 2011, Katz et al. [53] demonstrated that, as a result, forward wavefront shaping focuses the light both in space and in time. The authors applied the stepwise sequential algorithm for increasing the intensity of a two-photon signal. Even though they did not control the temporal degrees of freedom in the incident light, they observed that wavefront shaping resulted in a near-transform limited transmitted pulse. These findings nicely complement observations in acoustics, where it was shown that light can be focused spatially by controlling temporal degrees of freedom only [54, 55]. See Section 5.2 for a more detailed discussion of this phenomenon.

A second useful property of non-linear feedback is that optimization algorithms make the light converge to a single diffraction limited spot, even when the area where the feedback is coming from is larger than that. This ‘autofocusing’ property was demonstrated using the second harmonic signal from nanoparticles embedded in the scattering medium [56], two-photon fluorescence from a dye layer hidden behind a scattering medium [53], and non-linear photoacoustics [57] as non-linear feedback signals.

A non-linear feedback signal is proportional to $I^n$, where $I$ is the local intensity, and $n$ depends on the process that is used. If the feedback signal is coming from a 2-D source (e.g. a thin sheet of 2-photon dye), autofocusing will occur when $n \geq 1$. When the feedback comes from a homogeneous medium, a stricter condition of $n > 2$ is necessary [58]. The effect of varying $n$ was demonstrated in a beautiful experiment by Paudel et al. [59].

Because of autofocusing, the use of non-linear feedback is a very promising approach to generate a sharp focus with a ‘fuzzy’ feedback signal, thereby breaking through the limitations observed with linear feedback (see Section 3.4).

4.3. Photoacoustic feedback

When an absorbing structure is illuminated with a short laser pulse (usually several nanoseconds), the structure heats up rapidly. The resulting thermal expansion induces an acoustic wave that can be detected with an ultrasound detector, and subsequently be projected back to reconstruct the original structure. Imaging based on this photoacoustic (PA) effect is a rapidly finding biomedical applications [60]. PA imaging combines the specificity of optical absorption spectroscopy with the resolution of ultrasound imaging. Also, it provides a means of obtaining a feedback signal from deep inside a scattering medium without preparing the medium in any special way. The only requirement is that the medium contains absorbing structures.
The use of PA for wavefront shaping was pioneered by Kong et al. [61]. Chaigne et al. [62] extended the method to use the full 3-D photoacoustic reconstruction in order to measure a complete TM for at any arbitrary point in an absorbing structure. Imaging using PA-based feedback was demonstrated by scanning a sample through the focus [63].

Currently, the major limitation of PA based techniques is the large feedback area, which causes the contrast of the focus to be low (see Section 3.4). Demonstrations so far all used a geometry where the speckles in the medium were much larger than the optical wavelength, thereby artificially lowering the number of optical modes in the ultrasound focus. The use of non-linear photoacoustics [57] offers a promising approach to break through this limitation.

4.4. Other feedback

Any signal that is representative of the light intensity in a spatial or temporal target can be used as feedback. This flexibility allows for a large variety of target functions to be optimized. Apart from the feedback signals discussed above, other examples include the acousto-optic signal [64], and the polarization state of the output mode [65, 66]. A special case of iterative approaches was demonstrated by Nixon et al., who used an optical gain medium as an analog feedback mechanism to find lasing modes through scattering media [67]. Also, a promising direction may be to combine wavefront shaping with coherence gating, in order to achieve depth-selective focusing [68], or to reduce the amount of scattered light [69]. Feedback-based wavefront shaping can basically generate any spatio-temporal mode that one is able to detect; a feature that was convincingly demonstrated by Aulbach et al. [70] who used spatial wavefront shaping in combination with a scattering medium to guide light into any desired spatio-temporal mode.

5. Correlations and dynamics

So far, we have assumed that the medium is perfectly static, that perfectly monochromatic light is used, and that the incident wave has a perfect pointing stability. When one of these conditions is not met, the enhancement will decrease. Fortunately, this decrease is usually quite gradual. In this section, I will discuss what happens when one of these conditions is not met.

5.1. Dynamic media

If the medium is dynamic, its transmission matrix will change over time. Therefore, any feedback algorithm that aims to construct a matching wavefront needs to be fast enough to keep up with the dynamics of the medium, and continuously needs to keep updating the wavefront. If the feedback algorithm is slower than the medium dynamics, the constructed wavefront that is constructed has no relation with the current TM of the medium, and no enhancement will occur.

If a TM method is used, the medium should be stable over the duration of all \( N \) measurements. Since all measurements are performed sequentially, one should reduce \( N \) to match the decorrelation time of the medium [42]. As a rule of thumb, the maximum enhancement that can be achieved is \( \eta_{\text{max}} = \alpha T_m / T_p \), with \( T_m \) the time required for a single measurement, \( T_p \) the persistence time (speckle decorrelation time) of the medium, and \( \alpha \) was introduced in Section 2.1. Most iterative approaches automatically approach this maximum enhancement without adjustment to the sample correlation time [42, 50].

The advance of micromechanical light modulators has enabled wavefront shaping at a rate of over 20 kHz, which is 2000 times faster than the first wavefront shaping experiments (that used twisted nematic liquid crystal light modulators). The use of micromechanical light modulators for light focusing was pioneered by Conkey et al. using a digital micromirror device (DMD) [44], followed by Stockbridge et al. [71] who used a segmented mirror with a comparable speed. At such speeds, the PC that is traditionally used for executing the feedback
algorithm becomes the main bottleneck. Conkey et al. resolved this bottleneck by using programmable hardware for executing the algorithm and controlling the DMD; thereby eliminating the need for a PC in the feedback loop [72].

Once a matching wavefront is found, the enhancement will decay in time; exactly following the trend of the temporal speckle autocorrelation function [73, 74]. For long, it was believed that the relevant time scale for focusing light through the living skin was less than a millisecond (a typical time scale observed in dynamic speckle measurements). Surprisingly, in in-vivo experiments it was found that the reconstructed focus remains stable for a much longer time than that, up to several seconds [74, 75]. A possible explanation may be that there are two components to the speckle decorrelation: blood flow (fast), and dynamics of non-perfused tissue (slow).

The dynamics of living tissue is often considered a major hurdle for in-vivo applications of wavefront shaping. Fortunately, the future is bright: wavefront shaping can be done in several milliseconds, whereas a time scale of seconds is already fast enough for some in-vivo experiments.

Fig. 6. Temporal compression using only spatial wavefront shaping. a) In a conventional pulse shaper the pixels of the SLM are coupled to the spectral degrees of freedom by scattering from a grating; b) Similarly, coherent scattering in a random medium couples each SLM pixel to a different linear combination of the spectral degrees of freedom, forming a new random spectral basis that is phase-controlled by the SLM. In both cases, the spectral resolution maximum is determined by the optical path lengths differences in the medium/shaper. Reproduced from [53] with permission.

5.2. Broadband light

When the light is perfectly monochromatic, the optimal wavefront for focusing to a spot follows trivially from Eq. (1). For broadband light, however, the situation is slightly more complex since the temporal/spectral dimension needs to be taken into account too. Here, two regimes can be identified. If the bandwidth of the light source is smaller than the Thouless frequency $\omega$, the system can effectively be considered monochromatic [76, 77]. The Thouless frequency is inversely proportional to the Thouless time: the average time it takes light to diffuse through the sample. An equivalent criterion is to say that the correlation length
of the light source should be longer than the typical path length difference in the medium [78]. In the second regime, the bandwidth of the source exceeds the Thouless frequency. As a result, each spatial degree of freedom is associated with several independent spectral degrees of freedom [59]; increasing the dimensionality of the problem by one. In this regime, the scattering medium mixes spatial and temporal degrees of freedom, an effect well known from pioneering work in acoustics [54, 55]. As a result, spatio-temporal focusing can be achieved by only controlling spatial degrees of freedom [53, 56, 70] (see Fig. 6).

5.3. The memory effect

Under some conditions, the reconstructed focus is remarkably robust to a change in the angle of the incident beam. This feature is the direct result of the optical memory effect, a type of wave correlations that has been known since 1988 [79, 80]. The optical memory effect entails that a tilt in the incident wavefront results in an equal tilt in the transmitted wavefront. This way the transmitted field can be tilted without significant decorrelation; up to an angle known as the memory effect angle. The memory effect angle is proportional to \( \lambda / L \), with \( \lambda \) the wavelength and \( L \) the thickness of the sample [79–81]. When observed at a distance \( d \) away from the back surface of the sample, the transmitted pattern shifts over a distance proportional to \( d \lambda / L \).

The memory effect has been exploited for raster scanning fluorescence microscopy of objects hidden behind a scattering layer [21, 82]. Recently, it was discovered that the memory effect can also be used for computational imaging even without creating a focus in the object plane [83, 84].

The field of view of methods based on the memory effect is limited; the method can only be used for small objects placed at a large distance behind a thin sample (object size \( w < d \lambda / L \)). However, recently it was discovered that a different type of memory effect is present inside turbid materials that scatter light in the forward direction predominantly (such as biological tissue) [85]. Also, the memory effect angle for tissue was shown to be much larger than predicted [86]. These findings indicate that it may be possible to extend the concepts of memory effect imaging and scanning microscopy to work inside biological tissue.

6. Related topics

6.1. Adaptive optics

The term adaptive optics (AO) refers to a series of techniques to improve the image quality of an imaging system by correcting for aberrations in the optical path. Conventionally, AO works in the regime where the aberrations are smooth functions of the spatial coordinates, and where light can be described as paraxially propagating beams (e.g. [87]). As a result, most techniques for AO implicitly or explicitly use the smoothness of the aberration correction by expanding the wavefront in Zernike modes, or by using a deformable mirror with connected segments.

In contrast, wavefront shaping was conceived for the regime of strong multiple scattering. In this regime, the wavefront of the light is disordered on a length scale of \( \lambda / 2 \) [88]. Even more importantly, in AO each point at the deformable mirror can usually be associated with a single eikonal ray that propagates to the target. In wavefront shaping, each ray coming from the SLM results in a fully developed speckle pattern. Because this speckle pattern is the result of an infinite series of interfering, multiple scattered waves [29], it is not even possible to define a single optical path length for propagating from an SLM pixel to the target focus.
Table 2. Key differences between adaptive optics and wavefront shaping. Both techniques represent different ends of the scattering spectrum.

<table>
<thead>
<tr>
<th>Adaptive optics</th>
<th>Wavefront shaping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ray optics</td>
<td>Multiple scattering, diffraction, interference</td>
</tr>
<tr>
<td>Start with distorted image</td>
<td>Start with nothing (speckle)</td>
</tr>
<tr>
<td>Paraxial waves</td>
<td>Fully diffuse light</td>
</tr>
<tr>
<td>Mechanism: aberration correction</td>
<td>Mechanism: constructive interference</td>
</tr>
<tr>
<td>1 segment → 1 ray</td>
<td>1 segment → 1 multi-path interference pattern</td>
</tr>
<tr>
<td>Large field of view</td>
<td>Usually single point correction</td>
</tr>
<tr>
<td>Large bandwidth</td>
<td>Narrowband solution</td>
</tr>
</tbody>
</table>

The differences between the regimes of AO and wavefront shaping are summarized in Table 2. One should note, however, that these regimes just represent different ends of a continuum. Therefore, a strict separation cannot always be made. The region where the depth lies between one scattering mean free path and one transport mean free path can be considered a cross-over regime between these two fields. In biological tissue, this regime spans from $\ell_s = 100 \mu m$ down to $\ell_p = 1 mm$. In the last couple of years, ideas from AO and wavefront shaping are being combined to address specifically this biologically relevant regime (e.g [89]).

6.2. Phase conjugation

Optical phase conjugation is a well-established research field that started in the early 1970’s [90]. Like adaptive optics, most early applications were concerned with correcting smooth aberrations on a paraxially propagating beam [91, 92]. Interest in phase conjugation has renewed after an experiment by Yaqoob et al. [75], where it was shown that phase conjugation can also be used to focus light through biological tissue, for which light propagation is completely diffuse. A further breakthrough was the development of digital phase conjugation that allows digital processing of wavefronts, and permits an almost unlimited intensity gain [93, 94].

In the special case of paraxial propagation through an absorption-free medium, a large fraction of the source light can be collected by the phase conjugation system ($T_p = 1$ in Eq. (6)). Therefore, images of arbitrary complexity can be reconstructed with a high fidelity. For the case of biological tissue, however $T_p \ll 1$, which means that most of the light does not propagate back to the source. Still, the small fraction of light that propagates back to the original source is sufficient to create a focus with a high ($>10^3$) contrast [75].

Phase conjugation uses the time-reversal symmetry of light propagation to propagate scattered waves back to their original source. Since phase conjugation is essentially a single-shot process, it can be much faster than feedback-based methods. This speed comes at a price: light can only be focused back onto its original source. Although focusing light onto a light source may seem of limited use, the original source can be a nanoparticle that is easily embedded into tissue [36, 94], or a virtual light source created by acousto-optic tagging [12, 13, 15], offering very exciting possibilities for deep tissue imaging (see [27] for a review). The distinction between phase conjugation and iterative wavefront shaping is not always sharp: phase conjugation, too, can be applied iteratively [95, 96], in order to focus onto bright reflectors or to further shrink the generated focus [52].

6.3. Computational imaging

Since a ground breaking experiment by Bertolotti et al. [83], the fields of wavefront shaping and computational imaging have experienced a strong cross-fertilization. Bertolotti demonstrated that the optical memory effect can be used in combination with a phase retrieval algorithm to look through thin layers of scattering material. The exciting aspect of these
methods is that they do not require any a-priori information, and as such are truly able to see through opaque ‘walls’ [83, 84, 97].

A different class of computational methods requires the TM of the scattering sample to be measured in advance. Once the TM is known, it can be inverted to descramble any transmitted image [19, 22, 43, 47, 49]. Other promising computational approaches are under development, such as the use of spectral encoding [98]. Computational imaging methods ‘descramble’ a speckle pattern in order to look through a scattering medium. Even though no wavefront shaping is required for these methods to work, many of the concepts of wavefront shaping find a place in computational imaging, and vice versa.

7. Applications and outlook

In the last eight years, the field of wavefront shaping has expanded tremendously, and first applications are starting to be demonstrated. Clearly, one of the ‘holy grails’ of the field is to achieve microscopic-resolution imaging inside strongly scattering media. One immediately thinks of biomedical imaging, and how biology, medicine and neuroscience would change if humans could effectively be made transparent to visible light. Not only biomedical imaging would profit from such a technology: by definition all visible objects scatter light, so the range of possible applications is sheer unlimited.

In order to form an image, focusing light to a single point alone is not sufficient; typically the point needs to be scanned in a controlled manner to allow for e.g. raster scanning fluorescence imaging. When the scattering medium can be characterized in advance, TM methods can be used for scanning and imaging [16–19, 22, 43, 47, 49]. Even though the need for such a calibration represents a severe restriction, several groups have successfully implemented imaging through multi-mode fibers [16–19, 49, 99] based on these techniques, with applications in endoscopy. Also, scattering media may be used as microscope objectives for imaging at an extremely high resolution [21, 22]. In addition to TM methods, the optical memory effect can provide the ability to scan a focus in certain geometries [21, 82]. Especially the new anisotropic memory effect [85] opens new possibilities for deep tissue imaging. Finally, new feedback mechanisms are being explored that allow focusing light at arbitrary locations inside scattering tissue [57, 61, 62, 64]. Even though these methods are currently too slow for in-vivo imaging, there is still plenty of room for improvement both through engineering and through a better understanding of the physics of scattered light and sound/light interaction. These feedback-based methods may someday be combined with phase-conjugation-based techniques for deep tissue microscopy [12, 13, 15].

Imaging is certainly not the only application of wavefront shaping. The ability to manipulate light and to measure the TM of a sample has proven instrumental for fundamental scattering research [33, 39, 43]. In addition, wavefront shaping can be used for generating arbitrary spatio-temporal modes [70], high-resolution focusing [21, 22, 100], or for concentrating light inside nanoscale objects or plasmonic structures [101–103]. In addition, by shaping the incident light, disordered scattering materials can be ‘programmed’ to perform a large variety of optical functions, including beam splitters [104], spectrometers [105], polarization optics [65, 66], and even single-photon wavefront generators [106].

Striking examples of non-imaging applications of wavefront shaping are found in quantum-secure authentication [25], cryptographic key storage [26], and secure communication [24]. These applications disordered scattering materials as physical unclonable functions (PUF), since no existing method can duplicate a disordered scattering material in such a way that the TMs of the copy and original are identical [107].

In conclusion, the emerging field of wavefront shaping research is full of exciting new ideas with a huge range of potential applications. With creative ideas - combined with advances in engineering, electronics, and theoretical scattering physics - imaging through turbidity may just be one of many applications that the imaging community will be able to realize in the years to come.