Ultrasound-mediated optical tomography: a review of current methods

Daniel S. Elson, Rui Li, Christopher Dunsby, Robert Eckersley and Meng-Xing Tang

Department of Surgery and Cancer, Hamlyn Centre for Robotic Surgery, Department of Bioengineering, and Department of Physics, Imperial College London, Exhibition Road, London SW7 2AZ, UK

Imaging Sciences Department, Imperial College London, Du Cane Road, London W12 0HS, UK

Ultrasound-mediated optical tomography (UOT) is a hybrid technique that is able to combine the high penetration depth and high spatial resolution of ultrasound imaging to overcome the limits imposed by optical scattering for deep tissue optical sensing and imaging. It has been proposed as a method to detect blood concentrations, oxygenation and metabolism at depth in tissue for the detection of vascularized tumours or the presence of absorbing or scattering contrast agents. In this paper, the basic principles of the method are outlined and methods for simulating the UOT signal are described. The main detection methods are then summarized with a discussion of the advantages and disadvantages of each. The recent focus on increasing the weak UOT signal through the use of the acoustic radiation force is explained, together with a summary of our results showing sensitivity to the mechanical shear stiffness and optical absorption properties of tissue-mimicking phantoms.

Keywords: acousto-optics; ultrasound-mediated optical tomography; acoustic radiation force

1. INTRODUCTION

Optical imaging of tissue has great diagnostic promise owing to the strong and complex interaction of light with the different components of tissue, which has been shown to reveal rich physiological information. Besides the diversity of interaction methods and information available, light has the additional advantages of being non-ionizing and generally non-harmful within published average and peak intensity limits [1–3], and the possibility of applications across length scales from microscopic (micron or sub-cellular resolution) to macroscopic (tens of centimetres field of views). The main limiting factors that determine the design and performance of many tissue optical imaging systems are the strong absorption and scattering of light that occur in tissues. These limit the penetration depth of some optical imaging modalities to the superficial tissue layers up to a few millimetres. For instance, white light illumination, or narrow band reflection or diffuse imaging endoscopes, can image the surface layers of tissues in either a large field of view endoscope [4–6] or with microscopic resolution [7,8] at a few hundred microns imaging depth. To go beyond this depth and up to 1–2 mm, more sophisticated methods must be used such as two photon excited fluorescence, which uses the higher penetration depth of infrared light [9], or coherent techniques such as optical coherence tomography that detects only ballistic (reflected, back-scattered) photons using interferometry and rejects other scattered light [10].

For deeper tissue imaging, light that is in the ‘tissue optical transmission window’ (650–900 nm [11]) must be used, where absorption owing to haemoglobin is minimized and absorption owing to water is still relatively low. In general, it is not possible to detect ballistic photons, and the spatial resolution is degraded to approximately 10 mm by using techniques such as diffuse optical tomography (DOT) because of the many scattering events that occur as the light propagates through the tissue. The optical properties also limit the tissues that may be imaged to those that are relatively homogeneous and/or low absorbing such as...
breast and brain [12]. In DOT, the output light distribution for a number of different light inputs is recorded, which subsequently allows the distribution of scatterers and absorbers to be calculated, usually with only approximately 10 mm resolution at a few centimetres imaging depth owing to the ill-posed nature of the inverse problem [13]. It is normally assumed that the absorption mainly comes from oxy- and deoxy-haemoglobin, which have slightly different absorption spectra [11] and which may be used to obtain vascular and metabolic information about the tissue—important indicators of tumour development. A number of approximations may be made about the scattering properties of the tissue, so that the diffusion equation can be applied [12]. This has led to potential applications for DOT in the detection of breast cancer, in which higher concentrations of haemoglobin are used to identify cancerous lesions, and in neonatal brain imaging, in which the developmental response may be continuously and non-invasively monitored [12].

The aim of ultrasound-mediated optical tomography (UOT/UMT, often also referred to as ultrasound-modulated optical tomography, ultrasound-assisted optical tomography, acousto-photic imaging or acousto-optic imaging) is to achieve high spatial resolution sensitive to optical absorption by creating an artificial marker or beacon at some depth in the tissue. This beacon is created throughout the volume of the beam with highest spatial resolution and sensitivity at the ultrasound focal region. This ultrasound beam modulates the local refractive index and optical scatterer positions, which in turn modulates the phase of photons that pass through this region (sometimes these photons are referred to as ‘tagged photons’ in the literature [14]). These photons are highly scattered as they travel from the ultrasound volume to the boundary of the tissue, but the relative number of photons that passed through the ultrasound beam may still be approximated by recording the intensity of the signal that exits the tissue surface and is travelling in all directions owing to the high degree of optical scattering, leading to the requirement for large etendue (the product of the emitting surface and the solid angle) detectors. In addition, the random paths followed by photons reaching the detector result in random pathlength differences, causing an interference speckle pattern consisting of a distribution of bright and dark spots (a speckle pattern) [26]. The phase of the time-varying intensity modulation of these spots with respect to the ultrasound phase is random, and therefore simple spatial averaging of the speckle pattern by a single photodetector cannot directly increase the strength of the modulated optical signal. Furthermore, the random motions of the scatterers in the sample, for instance the movement of blood, decorrelate the speckle pattern and place a limit on the speed at which the signal must be collected. These effects will be described in more detail later in this paper.

UOT can be contrasted with photo-acoustic imaging (PAT) (also called opto-acoustic imaging), an alternative method using light and ultrasound that also combines the advantages of good ultrasound resolution with optical specificity. In this technique, short (nanosecond) pulses of light illuminate a tissue, causing a local thermoelastic expansion, which results in an acoustic wave. The distribution of optical absorbers in the tissue can be calculated by recording the temporal and spatial evolution of the ultrasound trace. Note that the resolution in PAT comes primarily from optical absorption contrast, whereas UOT may also be sensitive to changes in optical scattering [27]. For further information refer to recent reviews on this topic [22,28,29].

In this paper, we present a review of the mechanisms of ultrasound modulation of the optical signal, together with a description of the modelling and simulation methods that have been implemented to probe these mechanisms. We then outline the instruments that have been constructed to detect the ultrasound-modulated light. The potential for using the acoustic radiation force (ARF) to increase the UOT signal strength is introduced. We then present our recent results demonstrating the potential of the ARF for increasing the strength of the UOT signal and providing
additional contrast from the mechanical properties. Representative images of a scattering and absorbing phantom illustrate the improvement in signal strength and spatial resolution over the use of pure ultrasound.

2. THE ORIGIN OF THE ULTRASOUND-MEDIATED OPTICAL TOMOGRAPHY SIGNAL

Light is scattered multiple times during its passage through biological tissue, and a significant amount may be absorbed depending on the wavelength, the concentration of absorbers present (oxy- and deoxy-haemoglobin) and the thickness. If the sample is illuminated with a laser light source of sufficiently long coherence length then the intensity distribution at the output surface will be a laser speckle pattern consisting of a distribution of bright and dark ‘speckle grains’, owing to the coherent summation of the random phases carried by photons that have propagated along different highly scattered path lengths [26] (figure 1). When ultrasound is focused into the tissue it causes a periodic compression and rarefaction of the tissue, which has two main optical effects: a modification of the local refractive index, and a modification of the position of the optical scattering sites as illustrated in figure 1. Usually, ultrasound frequencies of approximately 1–5 MHz are used that provide a good compromise between axial resolution and penetration depth. Photons that travel through or near to the ultrasound focus then accumulate a phase modulation through these two effects as described in the following paragraphs.

For the displacement of scattering sites, a phase modulation of the transmitted light occurs owing to the change in the optical pathlengths that are modulated at the acoustic frequency [30,31]. This phase modulation results in the formation of acoustic sidebands either side of the optical frequency. Although the displacement of the scattering sites is extremely small—of the order of a few nanometres—it can be shown that strong scattering has a large effect on the output speckle field owing to the accumulated effect of multiple small contributions to the change in path of the photons caused by multiple scattering events [32]. Therefore, the nanometre ultrasound-induced displacements of the scatterers can have a large effect on the output. From a frequency-domain perspective, the electric fields of the photons scattered by the acoustically modulated scatterers become Doppler-shifted [33] (figure 1).

In addition to causing the displacement of scatterers, the acoustic field also causes density fluctuations in the medium, which result in a periodic variation of the refractive index. In the case of a homogeneous medium in the absence of scattering and absorption, a collimated laser beam that is perpendicular to the acoustic field will undergo Raman–Nath diffraction from the refractive index grating, and the diffracted beams will be frequency shifted by an amount equal to the acoustic frequency [34]. In a photonic picture of this interaction, the frequency of some of the transmitted photons is shifted by the frequency of the ultrasound wave through a photon–phonon interaction.

The ultrasound focal region

Figure 1. Schematic showing the propagation of representative highly scattered photons through a biological tissue in the presence of a focused ultrasound beam. The resulting speckle pattern at the output face is illustrated together with the modulation in intensity of a single speckle grain. Note that the scattering is low in this illustration, resulting in a well-defined optical beam, and that no absorbers are included.
biological tissue, the time-varying refractive index causes a periodic modulation in the optical pathlength, which introduces a frequency shift to an optical ray passing near the acoustic focus. Note that the stronger modulation that is possible by matching the Bragg diffraction condition is generally not possible for light propagating in tissue because the scattering mean free path is too small [25].

Both of these effects (refractive index and scatterer site modulation) result in the creation of a frequency sideband in the output spectrum separated from the original wavelength by the acoustic frequency. This frequency sideband is broadened owing to the other original wavelength by the acoustic frequency. This frequency sideband in the output spectrum separated from the site modulation) result in the creation of a frequency path is too small [25].

The shifting of the photon frequency by twice the frequency of the ultrasound has also been observed [30], and the interaction strength has been shown to vary as the square of the ultrasound amplitude [30,31].

### 3. MODELLING AND SIMULATION OF ULTRASOUND-MEDIATED OPTICAL TOMOGRAPHY SIGNALS

Analytical models of the effect of the mechanisms described in the previous section on UOT signals have previously been developed for different ultrasound conditions, including continuous wave ultrasound [36], pulsed ultrasound [37], non-uniform continuous ultrasound [38] and non-uniform pulsed ultrasound [39]. These models can provide insight concerning the effects of ultrasound parameters on modulated optical signals. However, the analytical models are limited by the assumptions used. For example, typically it is assumed that the ratio of the optical transport mean free path to the ultrasonic wavelength is large so that the ultrasound-induced optical phase variations associated with different scattering events are weakly correlated. This assumption is only approximately valid when the ultrasound frequency is higher than 2.5 MHz. Furthermore, it is assumed that the total phase variation owing to the ultrasound modulation is very small (much less than one radian). This condition is required so that the analytical autocorrelation function of the optical signal can be simplified, as described below. Although this may be a valid condition under normal ultrasound conditions [33], it may not be valid when high-amplitude ultrasound is used. A recent simulation study [40] suggested that the total optical phase variation could be above $2\pi$ with modest ultrasound amplitude, which could lead to a possible modulation ‘saturation’.

An alternative simulation method involves Monte Carlo modelling to calculate a series of photon paths through a stationary ultrasound field, which is essentially fixed in time compared with the time taken for the photons to propagate through the tissue (i.e. the simulation is carried out for one ultrasound phase) [41]. The additional optical phase accumulated along each photon path can then be calculated for the two UOT mechanisms, and this can then be repeated for different phases of the ultrasound cycle by adjusting the scattering particle positions and refractive index according to the strength of the acoustic field [41]. This code can also be implemented using a graphics processing unit to rapidly calculate whole speckle images [42]. Our simulation of the optical phase of the photons as they arrive at the output face of the sample using this method is shown in figure 2a. Note that this confirms that for moderate ultrasound pressures the optical phase modulation is low.

If a large number of photon paths are simulated then an autocorrelation can be calculated and the Wiener–Khinchin theorem applied to calculate the light intensity at the acoustic frequency. This method has been used to determine which of the above effects (scatterer displacement versus refractive index modification) is responsible for the majority of the modulated light under different acoustic conditions [41]. For a constant acoustic amplitude of particle displacement, the effect of increasing the size of the acoustic wavevector, $k_a$, is that the phase accumulated along the relatively longer optical transport mean free path is larger. Therefore, the phase modulation owing to the refractive index variation increases with $k_a$.

---

**Figure 2.** (a) Results from a simulation of a UOT system showing the phase shifts $\varphi$, for 9103 detected photons for a focal pressure of $P_{\text{focal}} = 2$ MPa and $k_a = 2.1 \times 10^4$. The histogram bins have a width of 2°. (b) The speckle pattern modulation depth contributed by the index of refraction alone, $M_n$; the modulation depth contributed by displacement alone, $M_d$ and the modulation depth contributed by both, $M_{\text{sum}}$; $k_a$ is the acoustic wavevector. The Monte Carlo simulation is based on Wang [41]. (a) Mean value, 6.7°; standard deviation, 10.8°; (b) blue circles, $M_{\text{sum}}$; red squares, $M_n$; green triangles, $M_d$. 

---

*R.S.S. Elson et al. (2011)*

[Image 166x673 to 297x778]

---

*Interface Focus (2011)*

Downloaded from http://rsfs.royalsocietypublishing.org/ on June 22, 2017
The contribution from scatterer displacement does not vary with $k$, because these displacements may be positive or negative and the sum after multiple random scatterers averages to zero regardless of the size of $k$. Simulation shows that (i) the contribution from the index of refraction is equal to the contribution from displacement when $k$ is less than a critical fraction of the transport mean free path ($0.6 \mu m$ from the analytical model [41]) and (ii) if $k$ is greater than the critical fraction of the transport mean free path, the contribution from the index of refraction is greater than the contribution from displacement. We have simulated this in figure 2b, obtaining results matched to those reported in Wang [41] and Elazar & Steshenko [43].

Yao & Wang [44,45] further extended their simulation of UOT to inhomogeneous scattering media with a confined volume of ultrasound. As mentioned in §1, the transmitted intensity of light through tissue is rapidly reduced as a function of scattering and absorption and tissue thickness. It has been shown that both the AC and DC components of the transmitted light are significantly attenuated as the tissue thickness is increased. However, the AC-to-DC ratio—i.e. the modulation depth—decreases less significantly [44]. Since the signal-to-noise ratio (SNR) is expected to be related to the modulation depth, it should be more sensitive to changes in optical properties than the optical transmittance, and implies that UOT can detect with higher absorption sensitivity than unmodulated diffuse intensity measurements.

Note that the phases of the photons exiting the tissue are also modified by other processes in biological tissues; for instance, the flow of blood, which consists of moving red blood cells that may scatter the light. This movement across a whole tissue, together with other diffusions and tissue movements, results in a randomization of the optical phase and causes the rapid decorrelation of the speckle pattern, characterized by a decorrelation time $\tau_c$. The decorrelation time is fast in tissue and also results in a corresponding broadening of the frequency spectrum [46]. The spectral broadening from human breast tissue has been measured to be a few kilohertz using a method of heterodyne detection similar to those described below [47], corresponding to $\tau_c$ of about 0.1 ms.

4. METHODS FOR DETECTING LIGHT MODULATED BY ULTRASOUND

4.1. Choice of experimental geometry and parameters

To achieve a good spatial resolution, focused ultrasound transducers can be used to spatially concentrate the ultrasound field. Although the acoustically modulated photons are generated with equal probability along the ultrasound axis, there is a relatively increased signal from optical absorbers within the ultrasound focal region. In order to improve the axial resolution, a number of further developments are required as described in §4.4. The ultrasound frequencies used are in the range of 1–5 MHz to provide a good tissue penetration and a small focal volume (approx. 300 $\mu m$ diameter, 9 mm depth of focus for 5 MHz ultrasound). There are many different geometries with which UOT signals may be detected, but in general, for convenience, the ultrasound and optical axes are often perpendicular to each other and the light is detected in transmission (e.g. [48]). There are also reflection geometries (e.g. [49–51]) and side-detection configurations [52] which may be advantageous for detecting signals from very thick tissues where transmission detection may not be possible. Granot et al. [53] have developed an analytical solution to express the strength of the signal and the SNR as a function of the ultrasound focal position.

When choosing a light source for UOT experiments, one of the main considerations is the requirement that a sufficiently long coherence length laser must be used with a high enough power to be able to yield enough photons transmitted (or reflected) from the sample. Monte Carlo modelling can again be used to simulate the propagation of many random photon paths through the tissue based on the simulated optical properties. Photons that are not absorbed in the simulated tissue can be traced to one of the output faces, allowing various useful parameters to be studied, for instance the pathlength distribution of the light [54]. For a 3 cm thick tissue, it is found that the range of pathlengths within 50 per cent of the mode pathlength is around 7 cm (i.e. the pathlength distribution spans approx. 3–10 cm in the tissue). Therefore, in order to maintain a clear speckle pattern on the output tissue face, the coherence length of the laser source should be approximately 7 cm or more [44].

This has led to the widespread use of 532 nm (frequency doubled Nd:YAG) lasers, although these are not optimal for obtaining good transmission, since the tissue optical transmission region is in the red and near infrared spectral region. Therefore, some groups have used single-mode laser diodes in the red [49] or near infrared [14], HeNe [50] and Ti:sapphire [55] lasers. More complex pulsed systems have been demonstrated, including a Nd:YAG ring oscillator amplifier [56] for increasing the optical detection efficiency. The light may illuminate a large region of the sample (approx. centimetres diameter) [14] or alternatively small single-fibre input may be used [49]. The choice of light source may also be determined by the operating wavelength of the optical detection device; for instance, the appropriate wavelength for a photorefractive crystal or a spectral filter. Multiple wavelengths have also been used to allow more quantitative tissue absorption measurements to be made [57,58].

4.2. Single-point detection methods

Early detection methods typically used a single-point detector that was either a similar size to one of the optical speckle grains or integrated over a number of these speckles. The two main detection methods may be classified as either heterodyne or interferometric. In heterodyne detection, a part of the illuminating laser is split and recombined with the signal beam at the detector, producing a temporal modulation as a result of the frequency-shifted signal beam. In the case of UOT, rather than splitting a part of the input beam, the

Interface Focus (2011)
heterodyne detection may take place between the light that was modulated at the ultrasound focus and that did not pass through the focus. The speckles that are formed from the scattering of the light in the sample are uncorrelated with each other, and so the largest modulation depth is obtained when the signal is collected from only one speckle grain, which severely limits the etendue. The second method (interferometric detection) combines the signal beam with itself using an optical filter or interferometer (Mach–Zehnder) [59]. It is not possible to separate the frequency-shifted light from the background unmodulated light using standard optical filters for UOT, but it is possible to use a Fabry–Perot interferometer to selectively detect the frequency-shifted light [60].

4.2.1. Fast photodetector and electronic filtering. The early reports on detecting the ultrasound-modulated light used single-point detectors [16,48,61] such as photodiodes or photomultiplier tubes (PMTs) that were able to detect both the unmodulated and the ultrasound-modulated light in a configuration similar to that shown in figure 3a. Electronic filtering could then be used to determine the amplitude of the AC

Figure 3. Summary of the main methods used to detect UOT signals. (a) Photodetector measurement, (b) confocal Fabry–Perot interferometer detection, (c) speckle image contrast analysis, (d) lock-in image detection, (e) heterodyne holographic detection, (f) detection using a photorefractive crystal, and (g) spectral hole-burning detection. PD, photodetector; AOM, acousto-optic modulator; PR, photorefractive crystal; \( n_O \), optical frequency; \( nUS \), ultrasound frequency; \( nX \), frequency shift (typically approx. 70–80 MHz), \( n_{CCD} \), charge coupled device (CCD) camera frequency.
component of the detected signal and thereby give an indication of the amount of light absorbed at the focus of the ultrasound transducer. An aperture was usually placed outside the sample and in front of the detector in order to control the size of the optical speckle so that it was approximately the same size as the detector. In a slight variation, a part of the incident laser light can be picked off before the sample and then recombined with the signal beam at the photodiode [31]. This laser light then provided a local oscillator and the signal was found by filtering the output from the photodiode at the ultrasound frequency.

With this fast detection method, the detection time can be within the speckle decorrelation limit, i.e. the total acquisition time must be much less than the time required for the speckle pattern to be affected by other processes such as the motion of blood cells or tissue motion [16]. Often, owing to the weakly modulated signals, a compromise must be made between an improvement of the SNR, through increasing the number of photons detected, and a reduction in the SNR, through speckle decorrelation [44].

4.2.2. Use of Fabry–Perot interferometer. Since the detection of the ultrasound-modulated light involves a frequency shift of the optical frequency by the ultrasound frequency, detection schemes have been proposed that can optically filter the unmodulated light so that the full dynamic range of the detector may be used. This overcomes the well-known problem of trying to detect a small signal on top of a large DC background carrying shot noise, i.e. it is better to detect the signal in a darkfield configuration.

One method to achieve this is to use interferometric detection with a confocal Fabry–Perot interferometer (CFPI), which is able to efficiently filter the background unmodulated light [59,62] (figure 3b). This idea, along with a number of others that are currently used in UOT, originally came from the detection of surface deformations as a result of acoustic modulation of the sample as a means of material inspection. Sakadžic et al. [60] used a high-frequency (15 MHz) high-resolution pulsed UOT system that incorporated a CFPI that was tuned to the laser frequency plus 15 MHz. This filtered out the background light achieving a measurement resolution of less than 100 μm at 3 mm tissue depth. The high intensity of the ultrasound pulses was able to overcome the noise limitation of the large bandwidth. CFPI can achieve a large etendue and is not affected by the decorrelation of the speckle (since they operate with incoherent light at fixed transmission frequencies). One limit is that the rejection of the background light is more efficient for higher ultrasound modulation frequencies and so the use of CFPI may be best suited to achieving high resolutions in thinner tissue samples. This has more recently been demonstrated in a technique for ultrasound-modulated optical microscopy, where acoustic frequencies of up to 75 MHz were used to achieve an axial and lateral resolution of 30 and 38 μm [63]. Such a high resolution (but at low imaging depth) has possible applications in detecting sentinel lymph nodes using an absorbing contrast agent [20] or for monitoring sub-surface vasculature [64].

4.2.3. Spectral hole burning. A technique for spectral filtering to leave only the light that has been frequency shifted to an acoustic sideband has recently been proposed by Li et al. [65] as an alternative method to the use of a confocal Fabry–Perot interferometer. The principle is that frequency-shifted light is transmitted by a spectral hole-burning crystal (a rare earth ion-doped optical absorber). This has a two-level atomic energy level structure and is inhomogeneously broadened, meaning that specific atoms have slightly different transition wavelengths. If it is cryogenically cooled, then the (homogeneous) broadening of these transitions is sub-megahertz [65]. When illuminated with light at an energy that is equal to one of the transitions, the crystal will absorb the light and subsequently reradiate it as fluorescence. If a sufficiently high-power laser is used, all of the atoms at a particular transition energy may become excited, causing saturation and relative transparency to additional photons at that wavelength.

For making UOT measurements, a high-power pump laser is frequency shifted to the ultrasound-modulated frequency sideband, and is used to burn a spectral hole in the crystal. The signal photons are then transmitted through the crystal while the unmodulated background light is absorbed (figure 3g). These crystals may be used at around 800 nm and are therefore well matched to the tissue optical transmission window. In the experiments reported, the pump light is pulsed by activating an acousto-optic modulator (AOM) so that a spectral hole may be burnt before the signal light arrives (the spectral hole filter lifetime is approximately 10 ms). Even though a speckle pattern is formed at the detector, the signal information is contained in the intensity of the light, and the detector may therefore integrate light from multiple speckles simultaneously and a high etendue may be achieved. The signal strength then increases with the square root of the number of speckles [26], and that speckle decorrelation does not affect the signal as the detection is incoherent. An axial scan may be achieved by resolving the photodetector intensity as a function of time during pulsed ultrasound propagation, and further work on the modelling of these crystals has been presented in Li et al. [66].

4.2.4. Parallel detection. The low light level present in a single speckle limits the detection time. Kempe et al. [31] calculated that the SNR would not be high enough for breast tumour diagnosis owing to the shot noise limit. Gross et al. [67] considered the efficiency of detection in terms of the optical etendue (the product of the solid angle of the light emission or collection with the active area) of the typical light collection schemes used. For the emitting surface of the sample, the etendue is πA, whereas, for a single speckle grain, the collection etendue is approximately A². Therefore, the collection efficiency for typical UOT systems was
calculated to be about $10^{-10}$ for a single pixel and $10^{-4}$ for parallel detection with $10^6$ pixels. The light collection efficiency can be improved by increasing the detection area, and the DC component will increase in direct proportion to this. However, since the phases of different speckles are uncorrelated, the increase in detected modulated photons will be countered by an averaging out of the modulation. The DC component increases more rapidly than the AC component as the detection area is increased, reducing the modulation depth by the square root of the number of speckles contained within the detection area [30,44]. In reality, there is a trade-off because neighbouring speckle grains are not necessarily completely uncorrelated, and there is also the noise characteristic of the detection system to take into account [31,61].

It is desirable therefore to simultaneously maximize the detection of the modulated signal by using multiple detectors that are of a similar size to the speckle grains, leading to the use of imaging detectors. However, the use of imaging technology imposes a maximum frame rate of hundreds to thousands of hertz on the detection—not fast enough to directly detect the megahertz optical modulation electronically—and various schemes have been proposed to detect the modulated light. These may be classified as:

- analysis of the spatial statistical properties of the speckle;
- parallel lock-in speckle detection; and
- heterodyne detection using digital holography.

These methods are described in the following sections, together with a technique that combines some of the principles of these methods—detection using photorefractive holography.

4.2.5. Laser speckle. The possibility to use properties of the optical speckle pattern that is formed when a laser (with a sufficiently long coherence length) is passed through an ultrasound-modulated scattering medium was proposed by some early patents in this field (e.g. [68]). Essentially, if the distribution of scatterers in the sample is stationary then the speckle interference pattern on the charge coupled device (CCD) is stationary. However, if the scatterers move on a timescale similar to the CCD exposure time, the camera records a moving speckle pattern, which causes the speckle grains to become blurred (figure 3c). The degree of blurring can be determined by measuring the contrast of the image [69], which is correlated with the number of photons that have been modulated by the ultrasound. The first demonstration of this technique in UOT was by Li et al. [70]. In these experiments, chicken breast of up to 25 mm thickness was imaged with gizzard inclusions. A low degree of speckle contrast was recorded (0.14) because no analysing polarizer was used and the speckle size was reduced to less than the size of 1 pixel to allow more light to be detected. This resulted in a low contrast difference for less absorbing inclusions.

This approach has an advantage over the other parallel detection methods in that only one image of the speckle pattern needs to be recorded and therefore the problem of sample movement and speckle decorrelation may be reduced, although the signal is reduced for long camera exposure times depending on the ratio of the decorrelation time to the exposure time. This method is the basis of the experiments reported later in this paper.

4.2.6. Parallel lock-in detection. An alternative approach uses parallel lock-in detection to allow the ultrasound-induced optical modulation to be determined for every pixel in a CCD array. The original scheme was proposed by Gleyzes et al. [71] and a system was constructed by Leveque-Fort et al. [14] that incorporated a $256 \times 256$ CCD array that was read at a frame rate of 50 Hz. The illuminating laser diode was modulated at the same frequency as the ultrasound frequency, and four 20 ms frames were read from the camera for different relative phases between the ultrasound and laser (figure 3d). From these images, the amplitude and phase of the modulated light could be calculated for each pixel. However, the signal is reduced over the period of the four-frame measurement owing to speckle decorrelation.

4.2.7. Heterodyne parallel speckle detection. As described above, one of the problems faced by methods to detect the modulated (signal) photons is that the signal is weak and sits on a large background. This leads to the requirement of large dynamic range detectors and even then the background noise will be large compared with the signal strength. An imaging method was proposed by Gross et al. [67] that is able to counter this problem through the use of heterodyne holography [72]. In this scheme, part of the laser is split from the main beam before it enters the sample in order to act as a reference local oscillator beam (as illustrated in figure 3e). If the local oscillator is then frequency shifted by the ultrasound frequency using an AOM (since the required shifts are small a pair of AOMs is usually used to shift up and then down by slightly different amounts) it can be recombined with the signal beam to produce interference between the local oscillator and the signal. An important principle of this technique is that the strength of the modulation depends both on the intensity of the signal and on the intensity of the local oscillator. Therefore by increasing the strength of the local oscillator, the amplitude of modulation may be increased, which carries information on the intensity of the signal. Therefore, the signal may be amplified above the camera read noise by the local oscillator without any background amplification and can become shot-noise limited [67].

To detect the modulation, a number of modifications are made to the schematic described above [67]. Specifically, the local oscillator frequency is shifted slightly higher than the acoustic frequency ($v_s + v_{CCD}/4$) to allow multiple images (typically four) of the interference pattern to be recorded on the CCD that have different phases between the signal and local oscillator. The modulation depth can then be calculated on a per pixel basis by using a simple calculation [73] based on the four input images (refer to figure 3e).
A second modification is that a narrow aperture is inserted close to the sample surface, which increases the size of the speckle grains produced from the self interference of the sample photons. Thirdly, the reference beam is incident on the CCD from an angle slightly offset from the signal beam. This shifts the spatial frequencies that are carrying the interference information between the reference beam and the signal beam to higher values and this can allow them to be distinguished from the low-frequency sample speckle. The speckle hologram may be Fourier transformed to allow the important spatial frequencies relating to the signal light to be selected and the speckle decorrelation noise to be filtered out [67]. The effect of speckle decorrelation during the acquisition time for this technique is therefore reduced both because of the shorter acquisition times achievable owing to the heterodyne gain and because of the reduction in the speckle decorrelation noise.

4.2.8. Photorefractive crystals. Again inspired by detection methods that are used in the optical study of acoustic surface effects for materials inspection [74], the use of photorefractive crystals was proposed [75,76]. This method allows the signal contribution for each speckle to be integrated on a single detector by using a holographic method that matches the reference beam wavefront to the complex speckled wavefront leaving the sample. The resulting increase in signal strength then allows a higher SNR to be achieved over single speckle measurements.

Photorefractive materials exhibit an optically induced change in refractive index (and/or absorption) in response to spatially non-uniform illumination. Fundamentally, the change in refractive index is due to an internally varying electric field that is created by photo-excitation of charge in response to the spatially varying incident illumination. The spatial variation in illumination gives rise to charge transport and trapping within the photorefractive medium. The resulting spatial modulation of the electric field is then converted into a change in refractive index through the electro-optical properties of the photorefractive, e.g. via the linear Pockels effect [77], creating a holographic refractive index pattern that can then cause diffraction of light. In the case where two beams are incident on the crystal (a signal and a reference beam), the refractive index grating can diffract some of the signal beam so that its wavefront is then matched with the reference beam. The result is that the random phases that are present in the speckle from the signal beam disappear and a low etendue single ‘pixel’ detector can then be used. The etendue of the whole detection system is then defined by the size and resolution of the photorefractive crystal employed. This combines the advantages of large area etendue detection without reducing the modulated signal by averaging out the modulation.

There are two main methods used for detecting the modulated light in practice. The first of these uses part of the illuminating beam as the reference beam [75], which results in a refractive index grating being written by the unmodulated photons. The reference beam is then diffracted into the signal beam path and is wavefront matched with the unmodulated light, which is then amplified at the detector. Although only the unmodulated light is amplified, the changes in the DC offset of this signal can be used to infer the amount of light that is modulated.

The second method (illustrated in figure 3f) uses a frequency-shifted reference beam so that only the small-modulated component writes the refractive index grating. The diffracted light from the reference beam is wavefront matched with the modulated light and a photodetector records the transmitted light with a heterodyned amplified modulated component (for more details on this mechanism see the heterodyne parallel detection section above). These two methods have been theoretically studied to determine the mechanisms of signal detection [33,78], and it has been suggested that the second method has advantages as it is effectively a background-free detection technique [35] although no clear advantage has yet been conclusively demonstrated.

There are a number of disadvantages with the photorefractive crystal methods: particularly, that the spectral response currently limits their use to longer infrared wavelengths, away from the region of low tissue absorption in the red and near infrared. However, recently a tellurium-doped tin thiophospholipidephosphate (SPS:Te) ferroelectric crystal was used at 790 nm, which holds promise for applying photorefractive crystals at the optical tissue window where water absorption is lower [79]. Also, the refractive index grating can only track the decorrelation of the speckle pattern caused by sample motions provided that the photorefractive response time is sufficiently fast [56]; otherwise, the signal is modified depending on the ratio of the decorrelation time to the photorefractive response time. It has recently been reported that a GaAs photorefractive crystal could achieve a response time of 0.25 ms, which is about the same value as the speckle decorrelation time for thick tissue [80]. Another disadvantage is that light from the strong reference beam may also be scattered and provide a high background light level at the detector [65].

4.3. Tissue phantoms

When phantoms are used, effort is made to try to match the optical absorption and scattering parameters as closely as possible to tissue. Many different types of simulated biological tissue have been used for UOT, including water with trypan blue (absorber) and poly-styrene beads (scatterer) [48], gelatin with trypan blue [61], agar with intralipid and Indian ink [81] or chicken and turkey breast tissue [14]. The first in vivo demonstration of UOT was by Lev & Sfez [46], who studied the light that was backscattered from a mouse and also a human forearm. This was collected by an optical fibre and detected by a PMT. This study illustrated that the speckle decorrelates much faster than in most tissue phantoms studied owing to the scattering by moving tissue components, for instance red blood cells flowing in the microvasculature, and the unrealistic speckle decorrelation times remain a limitation of phantom-based studies.
4.4. Methods of scanning the object

The most simple method to build up a three-dimensional map of the UOT signal is to physically scan the sample [48,61], or to move the ultrasound transducer [31]. In order to reduce the amount of scanning that is required to create two- and three-dimensional images of the sample and to provide axial (z-axis) resolution, Wang [23] used a frequency-swept ultrasound beam. By simultaneously modulating the PMT with a frequency-swept gain, different points along the ultrasound axis could be resolved owing to their different heterodyne frequencies in the output electronic signal. A Fourier transform of this waveform could then allow the ultrasound-modulated optical signal to be resolved into different frequencies corresponding to different positions along the ultrasound axis. This method also improved the spatial resolution because if the whole ultrasound column was to be simultaneously modulated at the same frequency then light interacting with the ultrasound anywhere inside the affected volume may be modulated [44].

The frequency-swept (chirped) approach also helps to improve the limited spatial resolution that can be achieved in the ultrasound propagation direction. Typical focused ultrasound transducers that are used in the megahertz frequency range at a few centimetres focal depths have focal regions that have resolutions of less than 1 mm laterally and 10 mm longitudinally [82]. The chirped pulse methods can allow the photons traversing different regions of this ultrasound focus to be resolved.

The chirped frequency scheme was initially only used with a single detector for detecting ballistic photons owing to the limited SNR that could be achieved, although it was subsequently expanded to exploit the parallel detection of multiple speckles using lock-in detection and a CCD (figure 3d). This was achieved by modulating the laser with a frequency-swept signal and simultaneously modulating the ultrasound with the same frequency-swept signal at a specific delay relative to the laser [83]. This delay time determined the spatial location along the ultrasound axis at which the lock-in method was effectively being applied, thereby achieving z-axis resolution. Chirped ultrasound and imaging [23,83] were subsequently combined by modulating the laser with a frequency-swept signal and simultaneously modulating the ultrasound with a frequency-swept signal at a constant frequency offset relative to the laser [84]. Heterodyne detection was then achieved at the CCD, and the signal from different spatial locations along the ultrasound axis could be found by calculating the fast Fourier transform on a pixel-by-pixel basis since the frequency spectrum simply encoded the z-axis location. One of the limits of this technique is that the scatterers in the sample must be stationary over the whole period of the chirp, i.e. there must be no decorrelation of the speckle pattern during the acquisition, which in this case was around 2 s. An alternative method that also provides axial resolution uses a random series of phase jumps in the ultrasound signal that is replicated and also applied to the optical signal. By varying the delay between the application of these two signals and using a photorefractive holographic readout scheme, a specific ultrasound axial position may be isolated [85].

The use of computer tomography methods has been proposed to further increase the optical resolution by overcoming the limit of the relatively long longitudinal ultrasound depth of focus (approx. 10 mm). The UOT signal detected can be considered as an integration of many incremental contributions from different points along the ultrasound axis. This is conceptually similar to CT imaging and, by rotating and translating the object and recording multiple projection angles, Radon transform methods could be used to make a reconstruction of the object [18].

Short ultrasound pulses can also provide z-axis resolution by recording the output AC optical signal with a high temporal resolution [50] as the pulse propagates. The use of pulses also allows for lower acoustic doses to be used that are in line with medical safety limits while maintaining the peak power, which is more standard for medical ultrasound systems. Since a high detector temporal resolution is required, this method cannot be easily incorporated into the imaging methods described previously, although a stroboscopic approach can be used. For instance, Atlan et al. [86] used the heterodyne holography method and showed that by stroboscopically activating the local oscillator they could resolve UOT data as the ultrasound pulse propagated through the sample.

An alternative method for obtaining resolution along the transducer propagation direction proposed by Selb et al. [82] used the detection of nonlinear effects as the incident ultrasound power was increased. The optical modulation at the second harmonic was studied; this is expected to occur where the peak powers are high owing to nonlinear processes, or from the interference of two waves that are both phase shifted at the ultrasound fundamental frequency, which causes interference at the second harmonic. It was suggested that this interference effect may be the dominant contribution owing to the experimental observation that the second harmonic optical modulation intensity increased with a quadratic dependence on the ultrasound power [82]. The signal detected at the second harmonic had a higher contrast and a better spatial resolution than the fundamental.

Spatial resolution can also be obtained by integrating UOT methods with standard ultrasound instruments. The efficient detection methods that are employed, using the photorefractive crystal approach in particular, have allowed images acquired with the standard ultrasound mode to be overlaid on images that were acquired using UOT so that automatically co-registered ultrasound/UOT images could be displayed [87].

There are a number of factors that cause spatial variation in the UOT signal that must be accounted for when scanning three-dimensional phantoms. When considering the standard UOT set-up, as depicted in figure 1, it can be shown using Monte Carlo modelling that, as the ultrasound focus is scanned along the y-axis from the light source towards the camera, the UOT signal strength decreases. This is due to the
higher relative numbers of photons that are modulated at the light input face that gives a higher signal-to-background ratio [44]. The ultrasound signal strength also decreases as the ultrasound probe is scanned in its axial direction as a result of the acoustic attenuation of the sample. It should also be noted that an increase in the spatial resolution may decrease the SNR of the already weak UOT signal.

5. ATTEMPTS TO INCREASE THE ULTRASOUND-MEDIATED OPTICAL TOMOGRAPHY SIGNAL USING THE ACOUSTIC RADIATION FORCE

The ultrasound displacement of scatterers is typically in the range of a few nanometres for clinically compatible ultrasound frequencies and intensities. The low optical modulation depth of the UOT techniques could potentially be increased by achieving larger scatterer displacements and therefore a greater modulation depth (until the signal saturates). One method of achieving this is to use the ARF instead of the pure ultrasound, which can cause tissue displacements of a few microns [88,89].

When ultrasound waves are attenuated, i.e. absorbed, reflected or scattered by a medium, a small and unidirectional force can be generated within the medium which is called the ARF. This is caused by the small momentum transfer from the ultrasound beam to the medium [90]. The amplitude of the ARF depends on the amplitude of the ultrasound, the attenuation coefficient of the medium and the cross-sectional area of the ultrasound beam. A common method to control the ARF is to alter the amplitude of the ultrasound; for example, a transient ARF can be generated by an ultrasound impulse, while an oscillatory ARF can be obtained by amplitude modulating the ultrasound. Another parameter that affects the ARF is the frequency of the ultrasound, since the tissue attenuation is roughly proportional to the ultrasound frequency. However, although the ARF in regions close to the transducer increases with higher ultrasound frequencies, the amplitude of ultrasound is more quickly attenuated and hence after a certain depth the ARF is reduced.

Although the first measurement of the ARF was made as early as 1903 [91], only recently has this physical phenomenon been widely explored in biomedical sensing and imaging applications. Nightingale et al. [92] were among the first to demonstrate the potential of measuring the ARF in vivo when they induced streaming within cysts to distinguish them from solid lesions. Sarvazyan et al. [88] demonstrated the potential of the ARF in imaging tissue elasticity properties, using the ARF to remotely induce shear acoustic waves that could then be measured to obtain tissue mechanical properties in a technique called shear wave elasticity imaging (SWEI). The measurement of the tissue displacement owing to the shear wave was made using an optical system and an MRI scanner, although it could be possible to use the same ultrasound transducer. In the same year, Fatemi & Greenleaf [93] proposed a similar technique to probe tissue properties using an oscillatory ARF generated by interfering two focused ultrasound beams (ultrasound-stimulated vibro-acoustic spectrography). The measurements were made with a separate acoustic sensor that was sensitive to the kilohertz frequencies of the ARF oscillation. Calcification on ex vivo human artery was clearly detected and visualized and the system demonstrated a reasonable spatial resolution of 700 μm. Nightingale et al. [92] reported a clinical feasibility study to image tissue elasticity of various tissues using transient ARF, where short (less than 1 ms) focused ultrasound was applied and the resultant tissue displacement was measured using an ultrasonic correlation-based method. It has been demonstrated that tissue displacements of the order of 10 μm can be generated by the ARF. More recently, the use of the ARF has been integrated into a clinical ultrasound scanner for ‘supersonic shear imaging’ that provides quantitative elasticity measurements that potentially could help in breast lesion characterization with B-mode ultrasound [94].

A few methods have been proposed to generate an ARF in UOT, including the use of short intense bursts of ultrasound [27,95–97] as well as amplitude-modulated (AM) ultrasound [40,55] and beating between two confocal ultrasound transducers [98]. These methods produce a time-varying intensity to the sample that causes a time-varying displacement of the scatterers. This displacement also depends on the acoustic absorption and reflection properties of the sample, since these both produce changes in momentum in the ultrasound wave, with high reflecting and absorbing samples producing a higher ARF. The final particle displacement, therefore, depends on the ultrasound strength, the local absorption or scattering and the local shear stress. An additional effect that results from the ARF is that shear waves are produced that propagate away in a direction perpendicular to the ultrasound direction [99]. The advantages and limits of using the ARF will be discussed in more detail in the final section below, summarizing our recent observations.

5.1. Use of the acoustic radiation force to detect optical and mechanical contrast

We have previously studied the trade-off in the use of the ARF between spatial resolution and signal strength when using an AM ultrasound trace to generate a periodically varying ARF in a tissue-mimicking phantom [100] using the set-up illustrated in figure 4. The phantoms used were made of agar with intralipid to provide an optical scattering coefficient of 5 cm⁻¹ and Indian ink to absorb the light. Although these phantoms did not have any specific ultrasound absorber or scatterer added, the high ultrasound frequency used (5 MHz) resulted in sufficient ultrasound absorption and a detectable ARF. The AM frequency was 500 Hz and CCD images were recorded with two different exposure times (0.2 and 2 ms). These were found to be sensitive either to pure ultrasound modulation (0.2 ms exposure), owing to the small distance that the optical scatterers moved during this low integration time, or to the ARF (2 ms exposure). A delay generator
was used to record speckle images of the scattered light field at different delay times with respect to the initiation of the ultrasound pulses (figure 4). Different inclusions with varying mechanical and optical absorption properties were inserted into the phantoms so that the sensitivity to these objects could be studied for the two regimes (0.2 and 2 ms exposure times). Illustrative results are presented in figures 5 and 6.

Figure 5. (a) Contrast difference versus CCD trigger delay time for a 250 Hz AM ultrasound burst with 0.2 ms and 2 ms CCD exposure time in a homogeneous area (blue line) and an absorbing area (red line). (b) Comparison between a one-dimensional profile of an optical absorber obtained with a 0.2 ms CCD exposure time measured 2 ms after launching a 250 Hz AM acoustic burst (red line) and one-dimensional profiles with a 2 ms CCD exposure time measured at varying delay times after launching a 250 Hz AM acoustic burst (adapted from [100]).
burst and at later times the propagation of the shear wave caused a blurring of this signal. The result for the 0.2 ms CCD exposure time is also included, showing a 40 per cent improvement in spatial resolution and an approximately 100 per cent improvement in image contrast when the ARF was used.

To demonstrate that the ARF may be used to detect the mechanical properties of the sample, a phantom was constructed that had an optical absorption and a mechanical discontinuity (made by adjusting the concentration of agar [81]), which had a lower shear stiffness, as illustrated in figure 6a. The two CCD exposure times (0.2 and 2 ms) were used to record speckle images, and the change in UOT contrast for a one-dimensional scan across the phantom is illustrated in figure 6. The red line was recorded for the 0.2 ms exposure time and shows that the standard ultrasound was able to detect the optical absorbing inclusion, but not the inclusion with a different shear stiffness. However, the 2 ms exposure time has a sensitivity to the ARF and also to any shear wave produced. The change in stiffness results in a modification to the shear wave produced and the system is, therefore, sensitive both to the optical and to the mechanical inclusion.

Two-dimensional images were recently measured with the same CCD exposure times (0.2 and 2 ms) and the same trigger delay times on a phantom containing an optical absorber (dimension 5 × 5 × 5 mm) by scanning the phantom. The results in figure 7a,b show two-dimensional images captured with a 0.2 ms CCD exposure time and a 2 ms CCD exposure time, respectively. The full-width half magnitude (FWHM) of the absorbing region was 6 mm for the ARF-sensitive 2 ms CCD exposure time and the FWHM was 7 mm for the traditional non-ARF signals measured with the 0.2 CCD exposure time. This represents an improvement of more than 17 per cent in spatial resolution. The differences in contrast between the maximum and the background were approximately 0.059 and 0.0423, respectively, which is a 40 per cent improvement in image contrast for the ARF-modulated signals compared with the pure ultrasound-modulated signals.

The spatial resolution and image contrast of the two-dimensional images does not increase as much as the one-dimensional result owing to the difference between the phantoms.

The spatial resolution and image contrast of the two-dimensional images measured with the short CCD exposure time (0.2 ms) do not change with time. However, when using the longer CCD exposure time (2 ms) and different CCD delay times (1, 5, 9 and 13 ms), the spatial resolution of the two-dimensional images decreases from a maximum at 1 ms. These results again demonstrate that short CCD exposure times tend to filter lower frequency vibrations from the ARF and are only sensitive to high-frequency movements at the ultrasound frequency. When using the longer CCD exposure time, the effects of low-frequency movements induced by the ARF and the resulting shear wave can be detected. With the optimal trigger delay time, the spatial resolution and image contrast can be improved simultaneously compared with pure ultrasound.

The increase in signal strength that can be obtained by using the ARF to make UOT measurements therefore has to be carefully balanced with the possible loss in spatial resolution owing to the propagation of the shear wave. The timing of the signal acquisition after the generation of the ARF should be kept as short as possible, and consideration should be made of the dependence of the attenuation of the shear wave as a function of ultrasound frequency. A potential additional benefit of the ARF is that it allows the simultaneous probing of the optical and mechanical properties of the tissue, since the signal strength is also affected by the stiffness of the tissue [81,100].

6. CONCLUSIONS

UOT has been shown to be sensitive to a number of tissue parameters, including the local absorption and scattering properties at the ultrasound focus, the ultrasound
absorption, the material shear stiffness and the shear wave attenuation. The measurement of these parameters either individually or in a combination would potentially provide a useful diagnostic signal for medical applications of this technique. The sophisticated detection techniques that have been developed to detect the weak UOT signals include the use of high-finesse interferometers, photorefractive crystals, spectral hole-burning crystals and heterodyne holographic techniques. These methods have demonstrated the potential for detecting sentinel lymph nodes, microvasculature, gold nanoparticles and for monitoring ultrasound therapy. The possibility of using multispectral methods to quantitatively and non-invasively detect these signals at depths of a few centimetres and with millimetre resolution has potential for breast cancer imaging owing to the possible sensitivity to tumour metabolism through the detection of oxygenation levels. At present the translation of this technology into a clinically relevant diagnostic device has been slow. The combination of better detection methods, higher UMT signals from the use of alternative ultrasound modalities and the possibility of probing the optical and mechanical properties simultaneously can be expected to open up further niche applications for these methods.

Figure 7. (a) A two-dimensional profile of an inhomogeneous phantom measured with a 0.2 ms CCD exposure time and a 2 ms CCD trigger delay time. (b) A two-dimensional profile of an inhomogeneous phantom measured with a 2 ms CCD exposure time and a 1 ms CCD trigger delay time. (c) One-dimensional profiles taking along the midline (Z = 9 mm) from (a) and (b). (d) One-dimensional profiles taking along the midline (Y = 0 mm) from (a) and (b). Circles with continuous line, CCD exposure time: 0.2 ms; squares with continuous line, CCD exposure time: 2 ms.

The authors would like to thank the UK EPSRC (grant nos. EP/H02316X/1 and EP/E06342X/1) and the Royal Society for their financial support. D.S.E. is supported by European Research Council Starting Investigator Award 242991.

REFERENCES


