Lock-in camera based heterodyne holography for ultrasound-modulated optical tomography inside dynamic scattering media

Yan Liu, Yuecheng Shen, Cheng Ma, Junhui Shi, and Lihong V. Wang

ARTICLES YOU MAY BE INTERESTED IN

High-speed single-shot optical focusing through dynamic scattering media with full-phase wavefront shaping
Applied Physics Letters 111, 221109 (2017); https://doi.org/10.1063/1.5009113

Focusing light through scattering media by polarization modulation based generalized digital optical phase conjugation
Applied Physics Letters 111, 201108 (2017); https://doi.org/10.1063/1.5005831

Pulsed ultrasound-modulated optical tomography using spectral-hole burning as a narrowband spectral filter
Applied Physics Letters 93, 011111 (2008); https://doi.org/10.1063/1.2952489
Lock-in camera based heterodyne holography for ultrasound-modulated optical tomography inside dynamic scattering media

Yan Liu, Yuecheng Shen, Cheng Ma, Junhui Shi, and Lihong V. Wang

Optical Imaging Laboratory, Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, USA

(Received 9 February 2016; accepted 28 May 2016; published online 8 June 2016)

Ultrasound-modulated optical tomography (UOT) images optical contrast deep inside scattering media. Heterodyne holography based UOT is a promising technique that uses a camera for parallel speckle detection. In previous works, the speed of data acquisition was limited by the low frame rates of conventional cameras. In addition, when the signal-to-background ratio was low, these cameras wasted most of their bits representing an informationless background, resulting in extremely low efficiencies in the use of bits. Here, using a lock-in camera, we increase the bit efficiency and reduce the data transfer load by digitizing only the signal after rejecting the background. Moreover, compared with the conventional four-frame based amplitude measurement method, our single-frame method is more immune to speckle decorrelation. Using lock-in camera based UOT with an integration time of 286 μs, we imaged an absorptive object buried inside a dynamic scattering medium exhibiting a speckle correlation time (τc) as short as 26 μs. Since our method can tolerate speckle decorrelation faster than that found in living biological tissue (τc ∼ 100–1000 μs), it is promising for in vivo deep tissue non-invasive imaging. Published by AIP Publishing.

[http://dx.doi.org/10.1063/1.4953630]

Light scattering prevents optical imaging from achieving high resolution deep inside scattering media. To break this optical diffusion limit (∼1 mm in tissue), ultrasound-modulated optical tomography (UOT),1,2 also called acousto-optic imaging,3 has been developed to image optical absorption4 and scattering5 properties with ultrasound-determined spatial resolution at depth up to several centimeters. Due to the acousto-optic effect, light passing through an ultrasonic beam undergoes a frequency shift to multiples of the ultrasonic frequency.6,7 By detecting the frequency-shifted light (tagged light), ultrasonically defined spatial resolution can be reached.

Various methods have been developed to detect the very few ultrasonically tagged photons out of a large background of untagged photons.23 Initially, a single-pixel detector, such as a photomultiplier tube4,7 or a photodiode,8 was employed. Since each speckle grain oscillates with a different phase, when N speckle grains fall on the detector, the useful AC signal amplitude is proportional to √N, while the informationless DC background is proportional to N.1,8 Thus, the largest modulation depth (defined as the ratio of the AC signal to the DC background) is obtained when the detector detects no more than one speckle grain, which severely limits the detection etendue9 (defined as the product of the detection area and the acceptance solid angle). To increase the detection etendue without reducing the modulation depth, three types of methods have been developed. The first type relies on a narrow spectral filter (∼MHz) to filter out the untagged light, so that a large-area single-pixel detector can be used. Examples include Fabry-Perot interferometers10–12 and spectral-hole burning13–15 based methods. These techniques are immune to speckle decorrelation due to motions of scatterers, but require bulky and expensive equipment. The second type of method uses crystal based holography to convert the wavefront of a reference beam to the complex wavefront of the sample beam, so the two beams can interfere constructively on a large single-pixel detector.16–19 This method, however, can be affected by speckle decorrelation due to scatterers’ motion inside living biological tissue, since the response time of the crystal is usually much longer than the speckle correlation time (∼0.1–1 ms) of living tissue.20–22 Recently, Ramaz’s group demonstrated the promise of SnP2S6:Te and Nd:YVO4 crystals for UOT, because of their short response times.23,24 However, these crystals usually work in a narrow range of wavelengths. The third type of method uses a pixel array, i.e., a camera, to detect the tagged light.9,25–28 Since the AC signal for each speckle grain is individually measured before being added together, this parallel method increases the modulation depth by √N compared with that of the original single-pixel detector based method, and achieves a much larger etendue. To increase the sensitivity, heterodyne holography9,29 was developed to boost the signal strength by a strong reference beam and thus achieved shot noise limited sensitivity.

In previous heterodyne holography based UOT,9 a camera recorded the beat formed by the tagged light and a planar reference beam at a frame rate of 4× the beat frequency. If averaging is not performed, four frames are required to reconstruct the amplitude of the tagged light. Thus, the time to obtain a UOT signal corresponding to one ultrasonic scanning position is limited by the low frame rates of conventional cameras (typically < 700 Hz with 300 × 300 pixels). Such a low speed makes this method inappropriate for in vivo applications since speckles decorrelate faster than 1 ms for thick living biological tissue (primarily due to blood
Moreover, conventional cameras record both the AC signal and the DC background for each pixel. When the modulation depth is low, which is the case in UOT since the tagged photons are buried in a large background of untagged photons, only a few bits of a pixel value can be used to represent the useful AC signal, while most of the bits are wasted in representing the informationless DC background, resulting in a low efficiency in the use of bits. For example, even with a 16-bit analog-to-digital converter (ADC), no more than 3 bits of a pixel value can be used to represent the signal, when the modulation depth is lower than $10^{-4}$ for each pixel. A modulation depth as low as $10^{-4}$ is realistic when UOT targets 10 mm deep inside chicken breast tissue, using a 50 MHz ultrasonic transducer (UT) for high spatial resolution.30 Besides the low bit efficiency, in previous methods all 16-bit data for each pixel, including both the small signal and the large background, were transferred to a computer. Since at least four frames of images must be transferred to calculate the UOT signal, the speed of data acquisition was severely limited by the heavy data transfer load.

In this work, we overcome the above drawbacks in heterodyne holography based UOT by using a lock-in camera31–33 (heliCam C3, Heliotis; 300 × 300 pixels, 40 μm pixel size), in which each pixel performs analog lock-in detection and outputs only the information of the AC signal, with a frame rate of up to 3800 frames per second to an on-chip memory. Specifically, the lock-in circuitry generates the in-phase ($S_\text{i}(r_i)$) and the quadrature ($S_\text{o}(r_i)$) components of the AC signal for each pixel with a position of $r_i$, $i = 1, 2, ..., 300 \times 300$, which are digitized by a 10-bit ADC. Then, using an on-chip field-programmable gate array (FPGA), the sought-after AC amplitude map $A(r_i)$ is calculated by $\sqrt{S_\text{i}^2(r_i) + S_\text{o}^2(r_i)}$ and transferred to a computer. Since the information of only the AC signal, not that of the DC background, is digitized, the lock-in camera dramatically increases the bit efficiency by using all the bits to represent the signal, and it enables the use of inexpensive low-resolution ADCs. Moreover, compared with conventional camera based heterodyne holography that needs to record and transfer 4 frames, our approach is able to obtain the AC amplitude map after only a single frame of measurement taken within 0.3 ms. It also reduces the amount of data to transfer by transmitting only one frame of the calculated AC amplitude map, instead of four frames of raw images composed of both the signal and the background. Moreover, the data processing on the computer is much simpler and faster, since we need only to add up the AC amplitudes of all pixels to obtain the UOT signal $S$: $S = \sum_{i=1}^{N} A(r_i)$, where $N = 300 \times 300$.

Morgan et al. developed a prototype sensor aimed for lock-in camera based UOT,32 and Barjean et al. recently developed a Fourier transform method to improve the axial resolution of this approach.34 However, both groups used custom-designed sensors with a relatively low etendue ($64 \times 64$ and $24 \times 24$ pixels), and so far few experiments in the UOT field have been done with dynamic samples which more closely resemble living biological tissue. Here, using a commercially available high-resolution lock-in camera, we performed UOT inside a dynamic scattering medium with a high-frequency ultrasonic transducer. We were able to acquire an image of an absorptive object (AO) with a lateral resolution of 153 μm, even when the speckle correlation time was as short as 26 μs, which is the shortest speckle correlation time ever reported in the UOT field.

The set-up for lock-in camera based UOT is schematically shown in Fig. 1. The output of a 532 nm continuous-wave laser (4 W, Verdi V5, Coherent) with a frequency of $f_0$ was split into a reference beam and a sample beam. After passing through acousto-optic modulators 1 (AOM1, AOM-505AF1, IntraAction) and 2 (AOM2, AOM-802AF1, IntraAction), the frequencies of the reference beam and the sample beam became $f_0 + f_a$ and $f_0 + f_{a2}$, where $f_a = 50$ MHz and $f_{a2} = 75$ MHz were the frequencies of the signals to drive the AOMs. The sample beam illuminated a dynamic scattering medium composed of two tissue-mimicking intralipid-gelatin phantoms35 (IP1 and IP2). IP1 (25 mm × 50 mm × 1.5 mm along the x-, y-, and z-directions) and IP2 (60 mm × 60 mm × 3 mm along the x-, y-, and z-directions) had a reduced scattering coefficient ($\mu_s'$) of 1 mm$^{-1}$. An absorptive object (AO, 2 mm × 1.6 mm × 1 mm along the x-, y-, and z-directions, absorption coefficient $= 5.4$ mm$^{-1}$) made of gelatin and black ink was attached to the surface of IP2, and the distance between the object and IP1 was 10 mm. During the integration time of the lock-in camera (286 μs, the minimum for our configuration), a 25 MHz spherically focused ultrasonic transducer (UT, V324-SM, Olympus), driven by a 200 Vpp sinusoidal wave with a frequency of $f_{a2} - f_a - f_b = 24.93$ MHz, emitted ultrasound along the x-direction to shift the frequency of a portion of the light passing through the ultrasonic focus to $f_0 + f_a + f_b$, where $f_b = 70$ kHz. These tagged photons beat with the reference beam at a frequency of $f_b$, and the lock-in camera was used to extract the AC amplitude of the beat (which was proportional to the amplitude of the tagged light) in each pixel with only a single frame of measurement. The reference beam and the sample beam (untagged light plus tagged light) on the camera had 24 mW and 53 μW of power, respectively. Since the shot noise was larger than the

![Fig. 1. Schematic of the set-up for lock-in camera based UOT inside a dynamic scattering medium.](image-url)
electron noise of the camera and the technical noise of the laser (see supplementary material), we achieved shot noise limited sensitivity.8 The ultrasonic transducer and the center of the object shared the same z position, and the position of the lock-in camera was adjusted so that roughly each pixel detected one speckle grain.

By translating IP1 along the y-axis at different speeds with a motorized stage, we were able to control the correlation time of speckles on the lock-in camera.22 To calibrate the relationship between the phantom movement speed and the speckle correlation time, we placed a conventional camera (pco.edge 5.5, PCO-TECH; pixel size = 6.5 μm) in the mirrored plane of the lock-in camera (see Fig. 1), to record movies of speckles when the phantom was moved at different speeds. A 6 × 6 pixel-binning was used to match the pixel sizes of the two cameras. Then, we calculated the correlation coefficients between the first and each of the ensuing frames of the recorded speckle patterns. By fitting the correlation coefficient versus time using22,36  

\[ y(t) = a \exp(-2t^2/\tau_c^2) + b \]

we obtained the speckle correlation time \( \tau_c \), defined as the time during which the correlation coefficient decreased to 1/e^2 at a given phantom movement speed. For example, Fig. 2(a) shows the correlation coefficient as a function of time when the phantom was moved at 0.01 mm/s, from which \( \tau_c = 59 \) ms was determined. The relationship between the speckle correlation time \( \tau_c \) and the phantom movement speed \( v \) is shown in Fig. 2(b). By fitting the experimental data with a theoretical model22, \( \tau_c = d/v \), we obtained \( \tau_c = 0.52v / \) [ms] (the unit of \( v \) is mm/s). Based on this equation, we were able to pre-set the speckle correlation time by controlling the phantom movement speed.

Figure 3 shows the UOT signal (\( S = \sum_{i=1}^M A(r_i) \)) and the root mean square (rms) UOT signal (\( S_{\text{rms}} = M \sqrt{\sum_{i=1}^M A^2(r_i)/M} = M \sqrt{\sum_{i=1}^M [S^2_i(r_i) + S^2_y(r_i)]/M} \)) as a function of the speckle correlation time \( \tau_c \) after self-normalization. When \( \tau_c > \tau_{\text{int}} \), the signal level is almost constant. In other words, speckle decorrelation does not affect the measurement, which is expected. When \( \tau_c < \tau_{\text{int}} \), however, the UOT signal decreases with decreasing \( \tau_c \). Surprisingly, the decay of the UOT signal is rather slow. Even when \( \tau_c = 26 \) μs, which is \( \tau_{\text{int}}/11 \), the signal decreases only to around half of the signal obtained when \( \tau_c = 100 \) ms.

To better understand the relationship between the UOT signal and the speckle correlation time, we developed the following analytical model. Let \( N \) denote the number of cycles integrated in the demodulation process of the lock-in camera. Then \( N = \tau_{\text{int}}/\tau_0 \), where \( \tau_0 \) is the period for one cycle at the lock-in frequency \( f_0 \), i.e., \( \tau_0 = 14.3 \mu s \). In our case, \( \tau_{\text{int}} = 20 \). Based on the principle of the lock-in camera, the in-phase signal \( S_1 \) for each pixel can be expressed as

\[
S_1(t_c) = \sum_{i=1}^N s_i(m\tau_0),
\]

where \( s_i(m\tau_0) \) is the in-phase signal for the \( m \)-th cycle in the beat. The autocorrelation function can be expressed as36

\[
g^{(1)}(\tau) = \langle s_i^*(t) s_i(t+\tau) \rangle / s^2_c \approx \exp[-(\tau/\tau_c)^2]
\]

where \( \langle \cdot \rangle \) denotes ensemble averaging over all camera pixels, \( \tau \) is a time delay, and \( s^2_c \) is a constant. After squaring both sides of Eq. (1) and taking the ensemble average over all camera pixels, we obtained

\[
\langle S^2_1(t_c) \rangle = N s^2_0 + 2s^2_c \sum_{m=1}^{N-1} (N-m) \exp[-(m\tau_0/\tau_c)^2].
\]

When \( \tau_c > \tau_{\text{int}} \), \( \exp[-(m\tau_0/\tau_c)^2] \approx 1 \), so \( \langle S^2_1(t_c) \rangle_{\tau_c \gg \tau_{\text{int}}} \approx N s^2_0 \). \( \langle S^2_1(t_c) \rangle \) can be derived similarly. Thus, the normalized rms UOT signal can be calculated by

\[
\frac{\langle S^2_1(t_c) \rangle}{\sqrt{\langle S^2_1(t_c) \rangle}}_{\tau_c \gg \tau_{\text{int}}} = \sqrt{1 + 2/N^2 \sum_{m=1}^{N-1} (N-m) \exp[-(m\tau_0/\tau_c)^2]},
\]

which is the analytical model used in Fig. 3. Although the theoretical

![FIG. 2. (a) The correlation coefficient between the speckle patterns as a function of time, when the phantom was moved at 0.01 mm/s. (b) The relationship between the speckle correlation time and the phantom movement speed. Errors bars are not plotted due to their indiscernible lengths in the figure.](image)

![FIG. 3. Normalized UOT signal as a function of speckle correlation time. Error bars are not plotted due to their indiscernible lengths.](image)

![FIG. 4. One-dimensional images of the embedded object acquired under different conditions. The data points denote the experimental data, which are fitted to a theoretical model (denoted by the lines). The signals are normalized by the maximum signal in each condition.](image)
predictions do not match the experimental data extremely well, the trends and the shapes are fairly similar.

To acquire an image of the absorptive object, we scanned the object along the y-axis and detected the UOT signal at each scanning position. During the whole process, IP1 was moved back and forth along the y-axis to cause the speckles on the lock-in camera to decorrelate. We took measurements only when the motorized stage reached a steady speed. Fig. 4 shows one-dimensional (1D) images of the object acquired when the speckle decorrelated at different rates. The dips in the images represent the object, which absorbed part of the tagged light. Although the signal-to-noise ratio is poorer when \( \tau_c = 26 \mu s \), the image qualities for \( \tau_c = 26 \mu s \) and \( \tau_c > 50 s \) (static phantom) are comparable in terms of image contrast, resolution, and the object’s width and position, demonstrating that we were able to acquire an image of the object even when \( \tau_c = 26 \mu s \). Fig. 4 also shows a direct transmission image in which the total transmitted sample light power measured by the conventional camera at each object scanning position is plotted. We could not observe the object in this image, since the resolution was too poor without doing UOT. To quantitatively determine the lateral image resolution \( r_{FWHM} \) of UOT (defined as the full-width at half maximum of the 1D point spread function), we fitted the experimental data with the following theoretical model: \( S(x) = \text{erf}(2\sqrt{\ln 2}(x-x_1)/r_{FWHM}) - \text{erf}(2\sqrt{\ln 2}(x-x_2)/r_{FWHM}) \), where \( a, d, x_1, x_2, \) and \( r_{FWHM} \) are fitting parameters, and \( \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp(-t^2) \, dt \) is the error function. From the fitting, the lateral resolutions are \( 247/\mu m, 286/\mu m, \) and \( 267/\mu m, \) when \( \tau_c > 50s, =1ms, \) and \( \approx 26\mu s \). These values are 1.6–1.9 times as large as the measured FWHM (=150/\mu m) of the 1D acoustic intensity profile, which is uncommon and possibly due to the absorption of the tagged photons by the object when they are tagged in the vicinity of the object.\(^{12,37} \) If the spans between 25% and 75% of the contrasts are used,\(^{19} \) the lateral resolutions are \( 142/\mu m, 163/\mu m, \) and \( 153/\mu m, \) when \( \tau_c > 50s, =1ms, \) and \( \approx 26\mu s \). The object width can be obtained by \( x_2 - x_1 \), which gives 1.51mm, 1.61mm, and 1.62mm when \( \tau_c > 50s, =1ms, \) and \( \approx 26\mu s \). These values are very close to the true 1.6mm width of the object.

In our set-up, there is a large gap (=10mm) between the phantom IP1 and the transducer UT, due to the walls of a water tank and a phantom mount. Because of this distance, after passing through IP1, the light diffused to a big blob on the x-y plane where the transducer resides, even though IP1 was not thick compared with the samples used in previous studies. The modulation depth for each camera pixel was measured to be 0.12%, which is relatively low among the numbers reported in the literature.

Currently, even though the measurement is done within 0.3 ms, it takes at least 6 ms to transfer the data from the lock-in camera to the computer, limited by the low speed of the USB 2.0 interface (~250 MB/s). If a camera link interface is used (data transfer rate ~7024 MB/s), the data transfer time can be reduced to ~0.2 ms. Another way to lower the transfer time is to reduce the data load. Instead of transferring an AC amplitude map, it would be better to calculate the summed AC amplitude (\( S = \sum_{i=1}^{M} A(r_i) \)) using the on-chip FPGA,\(^{34} \) and transfer only this number instead of one frame of 300 \( \times \) 300 numbers. This simplification can greatly reduce the data load and thus the data transfer time.

The axial resolution was poor in our experiment, since a long burst of ultrasound was used to make sure the beat signal existed during the integration time of the lock-in camera (286 \( \mu s \)). To further improve the axial resolution, a frequency-swept method\(^{38} \) or a Fourier transform method\(^{44} \) can be employed.

To conclude, in this work, we used lock-in camera based UOT to acquire high-resolution images of an absorptive object inside a dynamic scattering medium. Our method is bit efficient. It can finish the measurement within 286 \( \mu s \), which is comparable with the typical speckle correlation time of living biological tissue (100–1000 \( \mu s \)), and it can tolerate speckle correlation time as short as 26 \( \mu s \). With these capabilities, our method holds promise for in vivo deep tissue non-invasive optical imaging.

See supplementary material for a noise analysis.

This work was sponsored in part by National Institutes of Health Grant Nos. DP1 EB016986 and RO1 CA186567.


