Published in final edited form as:

Ultrasonics. 2015 February ; 0: 456–472. doi:10.1016/j.ultras.2014.09.014.

Complex Background Suppression for Vibro-acoustography Images

Matthew W. Urban^{*,1}, Chenyi Wang², Azra Alizad¹, and Mostafa Fatemi¹

¹Department of Physiology and Biomedical Engineering Mayo Clinic College of Medicine Rochester, MN 55905

²Department of Bioengineering University of Washington Seattle, WA 98105

Abstract

Vibro-acoustography (VA) is an ultrasound-based imaging modality that maps the acoustic response, or acoustic emission, of an object stimulated by two ultrasound waves at slightly different frequencies. VA images typically have a nonzero background intensity which can reduce contrast in images. We present a method that uses the complex representation of the acoustic emission data to estimate and suppress the unwanted background signal. This method utilizes a fast, linear approach to the problem called complex background suppression (CBS) using a square filtering window of size $W \times W$. Images processed with the CBS algorithm have significantly enhanced contrast. Another improvement observed with this method is the ability to better localize objects within the depth direction with respect to the ultrasound transducer. This algorithm was tested on images obtained from scanning a phantom with spherical inclusions, a urethane breast phantom, and *in vivo* human breast. The results show that image quality is improved through processing with the CBS algorithm by increasing the contrast of features in the images. The contrast in the sphere phantom was increased by factors of 2-12 depending on the sphere. Utilizing the CBS algorithm increased the contrast in breast phantom by factors ranging from 1.1-5.4 for various inclusions. The size of the filtering window, W , affected the contrast achieved between the phantom features such as the spheres or simulated inclusions and the background material. Application of the CBS algorithm also demonstrated that objects could be localized in depth much better as the relationship to image intensity level was directly correlated to objects located at the center of the focal plane in the axial direction. This method has wide applicability for all VA imaging applications.

Keywords

vibro-acoustography; complex; ultrasound; parametric mixing; background artifact; acoustic emission

© 2014 Elsevier B.V. All rights reserved.

*Corresponding author urban.matthew@mayo.edu Phone: (507) 538-1522.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

C. Wang contributed to the paper while he was a visiting student at the Mayo Clinic College of Medicine during the summer of 2011.

1. Introduction

Vibro-acoustography (VA) is an imaging modality that uses two beams of ultrasound and co-focuses these beams in a small region to produce an acoustic response at the difference frequency (Δf) between the two ultrasound beams [1, 2]. The interaction of the beams with an object creates a sound field called acoustic emission at Δf which is measured by a nearby hydrophone. The co-focus of the two beams is mechanically or electronically translated over the object in a raster-scan pattern to create an image. The ultrasound frequencies typically used are 3-5 MHz which are in the same range used for medical ultrasound imaging, and Δf varies over a range of 10-70 kHz for most applications.

Vibro-acoustography has been used in imaging of arteries [3], breast [4, 5], thyroid [6], prostate [7-9], and in other biomedical applications [10]. It has also been used in nondestructive evaluation and imaging of surface roughness [11, 12]. A complete review of the applications of this method was detailed by Urban, *et al* [13].

VA images have unique characteristics. First, VA images are not degraded by the speckle artifact found in conventional ultrasound B-mode images. Even though ultrasound is used to produce the excitation in a very small area by virtue of ultrasound beam focusing, the content of the image is directly related to the value of Δf which has long wavelengths in tissue (2-15 cm), so the coherent interference present in conventional ultrasound B-mode that gives rise to speckle is mostly avoided.

Another characteristic of VA imaging is that the point-spread function (PSF), produced by using a spherically focused ultrasound transducer with two elements enhances the edges of scanned objects. The PSF acts similar to a differentiating operator which enhances the appearance of edges [10]. This is particularly useful to delineate structures such as calcifications and lesions.

Another aspect of VA images that has been observed is that soft tissue features can be depicted, but depending on the object and system setup, the contrast between these features and normal background tissue can be rather low. Even when there is no object or ultrasound scatterers present, an acoustic emission signal is measured. This is typical when we perform experiments in a water tank, and we have observed signals when the transducer focal region is located in the water or pure gelatin with no scatterers. Ideally, we would expect minimal or no signal from these locations because of the absence of an object or ultrasound scatterers.

In this paper we will assume that they are steady-state sinusoidal signals that can be characterized by their magnitude and phase. In most cases, VA images are created from the magnitude of the signals, and the phase data is largely ignored. A few studies have shown that the phase images have interesting contrast which can depict features better than the magnitude can including brachytherapy seeds implanted in an *ex vivo* prostate [1, 2, 8]. The magnitude and phase data can be represented as a single complex quantity. It is the aim of this paper to take advantage of the complex data to eliminate background effects to improve contrast in VA images.

The paper will be organized as follows. We will describe characteristics of typical VA images and explain why contrast can be low. A filtering scheme will be proposed that uses both the magnitude and phase information of the signals to improve the contrast by suppressing the signals that occur in the background regions. This filtering scheme was testing in two different phantoms and two *in vivo* breast imaging data sets. We will present the results from application of the filter and discuss the results.

2. Methods

The signal received at the phase sensitive hydrophone is the summation of signals produced by different physical phenomena. These mechanisms that can create acoustic signals at f include dynamic acoustic radiation force [2, 13-15], parametric mixing [16-19], nonlinear scattering from the object [15, 18, 20], and transient vibration from the transducer. We consider all of these mechanisms in the context of two incident propagating waves at different frequencies.

2.1 Mathematical Model

For the general mathematical model of VA acoustic signal sources, we will consider the four primary sources discussed above. We will combine the contributions from these sources using a linear summation.

The radiation force can occur where there is a medium that is either acoustically reflective or absorptive at the focus or along the propagation path, and the acoustic signal arising from this component is denoted as $s_{RF}(\mathbf{x}, t)$. In the subsequent analysis we will not distinguish whether the radiation force is being generated because of reflection or absorption, but note that both may be present.

Although, not addressed in depth by previous publications, dynamic radiation force can occur on the path to the focus wherever two ultrasound beams spatially overlap and there is either absorption or reflection. This mechanism has largely been ignored in VA imaging, because it has been assumed to be small compared to the force present in the focal region, but it should not be totally discounted. Depending on the spatial distribution of the two beams that are used, this component could be non-negligible. Figures 1(a)-(b) shows the simulated pressure fields for each of the elements of the confocal transducer that was used in the experiments that will be described later. Each image is displayed with a dynamic range of 50 dB. The custom-made transducer built in-house has two elements, a central disc and an annular ring which have a center frequency of 3.00 MHz. The central disc has an outer radius of 14.8 mm. The annular ring has an inner radius of 16.8 mm and an outer radius of 22.5 mm. The elements are spherically focused at a depth of 70 mm. The pressure fields were simulated using Field II [21, 22] using frequencies of $f_1 = 3.00$ MHz and $f_2 = 2.95$ MHz so $f = 50$ kHz. The magnitude of the pressure of each of the pressure fields generated at f_1 and f_2 were recorded at each location. To extract the f component of the dynamic force at each spatial location as shown in Fig. 1(c), the pressure signals were summed, and the sum was squared. A fast Fourier transform was computed and the component at f was recorded.

The point-spread function (PSF) for VA is directly related to the volume over which the radiation stress is generated, and this volume is determined by the region where the two primary beams overlap. As can be seen in Fig. 1, the fields at f_1 and f_2 are concentrated near the focus at 70 mm. Dynamic radiation force can be produced where both beams overlap and cause a temporally modulated acoustic pressure at f . In the case of these two beams transmitted from this confocal transducer, the overlap is not significant until the beams reach the focal area. However, the overlap region has a long depth-of-field so there could be appreciable effects from temporally modulated pressure which is time-averaged to produce the dynamic radiation force in the nearly 1 cm length encompassed by the overlap region. Therefore, the overall volume interrogated by the scanning process is the product of the scan area and the axial extent of the PSF in Fig. 1(c). However, the magnitude in the axial direction would be weighted so some details that are not at the center of the focal plane would not contribute as much to the measured acoustic signal. The effects of depth-of-field of the PSF on image formation for VA was dealt with by Silva, *et al* [23].

The parametric mixing component, $s_{PM}(\mathbf{x}, t)$, can occur along the path and at the focus if there is no object that presents significant acoustic reflection or absorption present. As with the radiation force the parametric mixing at the focus will be much larger than that generated along the path for the case of the confocal dual beam configuration because, as stated earlier, the overlap region of the two primary beams is centralized about the focal plane of the transducer.

The nonlinear scattering component requires the presence of an object to scatter the incident waves, and the component is denoted by $s_{NS}(\mathbf{x}, t)$. Lastly, we denote the acoustic signal from transient vibration of the transducer is denoted by $s_V(\mathbf{x}, t)$.

The sum of the primary sources can be written as

$$s(\mathbf{x}, t) = s_{RF}(\mathbf{x}, t) + s_{PM}(\mathbf{x}, t) + s_{NS}(\mathbf{x}, t) + s_V(\mathbf{x}, t), \quad (1)$$

and a schematic of this summation is shown in Fig. 2.

For modeling purposes, we will consider that all these primary sources can be modeled as the small independent disks as described in the development by Fatemi and Greenleaf [2]. The typical values of f range from 30-70 kHz, and the disks we consider are on the order of a millimeter (the width of the PSF main lobe) so the acoustic waves can be considered non-directional spherically propagating waves.

We can model the signal received at the hydrophone by considering that the primary sources each have magnitude and phase evaluated at ω that will depend on the position of the transducer with respect to the object and hydrophone. We apply a Fourier transform to examine the magnitude and phase for each of the primary sources.

$$\begin{aligned}
S(\mathbf{x}, \Delta\omega) = & |S_{RF}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{RF}(\mathbf{x}, \Delta\omega) + \Delta\omega l_{RF}/c)} \\
& + |S_{PM}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{PM}(\mathbf{x}, \Delta\omega) + \Delta\omega l_{PM}/c)} \\
& + |S_{NS}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{NS}(\mathbf{x}, \Delta\omega) + \Delta\omega l_{NS}/c)} \\
& + |S_V(\mathbf{x}, \Delta\omega)| e^{-i(\theta_V(\mathbf{x}, \Delta\omega) + \Delta\omega l_V/c)},
\end{aligned} \tag{2}$$

where S , S_{RF} , S_{PM} , S_{NS} , and S_V are the Fourier transforms of s , s_{RF} , s_{PM} , s_{NS} , and s_V , $|\cdot|$ is the magnitude operator, θ is the phase of the acoustic wave, l is the distance the acoustic wave travels from the primary source at $\mathbf{x}$ to the hydrophone at $\mathbf{x}_H$ such that $l_n = |\mathbf{x} - \mathbf{x}_H|$.

We can rewrite (2) using the relationship $\omega/c = 2\pi/\lambda$ as

$$\begin{aligned}
S(\mathbf{x}, \Delta\omega) = & |S_{RF}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{RF}(\mathbf{x}, \Delta\omega) + 2\pi l_{RF}/\lambda)} \\
& + |S_{PM}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{PM}(\mathbf{x}, \Delta\omega) + 2\pi l_{PM}/\lambda)} \\
& + |S_{NS}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{NS}(\mathbf{x}, \Delta\omega) + 2\pi l_{NS}/\lambda)} \\
& + |S_V(\mathbf{x}, \Delta\omega)| e^{-i(\theta_V(\mathbf{x}, \Delta\omega) + 2\pi l_V/\lambda)}.
\end{aligned} \tag{3}$$

The usefulness of this form is that the phase term ($\psi = 2\pi l/\lambda$) in the summation terms has the propagation path scaled by the spatial wavelength. This will be useful when examining the background suppression method described later. The other noteworthy element is that the phase θ is determined by the PSF of the particular source, and the phase ψ is determined by the path length that the acoustic wave travels from the source to the hydrophone.

We can simplify this general model with a few practical considerations. From our experience with VA scanning we have found that by employing temporal shading of the toneburst signals before applying them to the piezoelectric elements the acoustic transient vibration is greatly diminished and the contribution from S_V approaches zero [24]. Holding with this consideration (3) becomes

$$\begin{aligned}
S(\mathbf{x}, \Delta\omega) = & |S_{RF}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{RF}(\mathbf{x}, \Delta\omega) + 2\pi l_{RF}/\lambda)} \\
& + |S_{PM}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{PM}(\mathbf{x}, \Delta\omega) + 2\pi l_{PM}/\lambda)} \\
& + |S_{NS}(\mathbf{x}, \Delta\omega)| e^{-i(\theta_{NS}(\mathbf{x}, \Delta\omega) + 2\pi l_{NS}/\lambda)}.
\end{aligned} \tag{4}$$

In summary, at locations where there is an object or scatterers to reflect, absorb, or scatter the ultrasound, some signal at f will occur either due to the radiation force and nonlinear scattering mechanisms. At locations where objects or scatterers are absent, then the parametric mixing mechanism will be dominant because the beams will still intersect, but no radiation force or scattering will occur. In the case of soft tissue, there may be contributions from all three primary sources, but one is typically dominant over the others.

2.2 VA Image Formation and Characteristics

2.2.1 Image formation model—We turn now to the modeling the image formation from the acoustic fields produced by the different signal sources described above. The acoustic

pressure signal $S(\mathbf{x}, \omega)$ measured at the hydrophone is measured when the excitation is focused at $\mathbf{x}$. This is typically performed in a raster scanning scheme. VA is fundamentally a C-scan ultrasound modality, which means that the information contained in the image is localized to the focal plane that is a constant distance from the transducer surface represented by z_F . Multiple images at different imaging planes are typically taken to obtain data from a three-dimensional volume.

2.2.2 VA scanning—At each focal location tonebursts with typical lengths of 100-150 μs are transmitted, the waves interact with the object and medium, and the resulting acoustic pressure is recorded. In most cases, hardware signal conditioning is performed by applying a bandpass filter (30-80 kHz bandwidth assuming $f = 50$ kHz) and some amplification is applied. The signal is digitized for further software post-processing. In software, a bandpass filter that is more narrowband and centered about the value of f is applied. The signal is temporally windowed, and the length of the window is selected based on the length of the transmitted toneburst.

For a given imaging plane at z_F , a fast Fourier transform of the time-domain signal at (x, y) is applied to $s(x, y, t)$ to obtain $S(x, y, \omega)$ and images of the magnitude, $M(x, y)$, and phase, $\phi(x, y)$, are formed by extracting the data at frequency ω for each pixel where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ are the real and imaginary parts of the signals.

$$M(x, y, \Delta\omega) = \sqrt{\text{Re}(S(x, y, \Delta\omega))^2 + \text{Im}(S(x, y, \Delta\omega))^2} \quad (5)$$

$$\phi(x, y, \Delta\omega) = \tan^{-1} \left(\frac{\text{Im}(S(x, y, \Delta\omega))}{\text{Re}(S(x, y, \Delta\omega))} \right) \quad (6)$$

It should be noted that the tonebursts at the two different ultrasound frequencies are transmitted with the same relative initial phase difference and digitizing of the acoustic emission is synchronized with the transmission of the tonebursts. The signal is temporally gated such that a continuous wave can be assumed for the magnitude and phase calculations. These conditions provide a constant reference for the phase measurements.

2.2.3 VA image characteristics—The resolution of VA images has been extensively characterized by analyzing the PSF of the excitation configuration for many different types of configurations of the ultrasound beams [2, 14, 23, 25-29]. The contrast in VA images has not received the same level of attention, but this concept is critical to distinguishing diseased tissue from normal tissue. The contrast in VA images is determined by a few different factors related to parameters of the object related to interactions of the object with the ultrasound waves and the acoustic response. The acoustic impedance or ultrasound attenuation impacts the level of radiation force or nonlinear scattering produced. The acoustic response is also partially determined by the frequency response of the object at f . Additionally, the nonlinearity of the medium has an impact on the acoustic signal being produced by parametric mixing.

The ideal VA image would depict features that are unlike the normal background tissue with high contrast. For example a phantom with spherical inclusions embedded in a gelatin block with no ultrasound scatterers should ideally have a black background and the spherical inclusions should appear with high image intensity. However, in practice VA images of such phantoms visualize the spherical inclusions, but the images also have a uniform gray background where there are no ultrasound scatterers or reflective objects [2, 14, 28]. This inherent background signal, which could originate from a number of sources, including parametric mixing, reduces the overall image contrast in a majority of VA images. It would be advantageous to have a method to remove this background signal from VA images to improve image contrast.

2.3 Background Suppression

2.3.1 VA Background Characteristics—The nonzero background signal in VA images is typically fairly uniform with variations that have very low spatial frequency. The wavelength of acoustic emission at $f = 50$ kHz is 30.8 mm, assuming the speed of sound in the medium as 1540 m/s, and the images that are acquired are typically 50×50 mm so only a few wavelengths of the acoustic emission are contained within an image. This type of variation was examined by Urban, *et al.*, by imaging large uniform gelatin blocks [24]. The spatial variations in these images changed based on the value of f . These variations can be explained by examining Eq. (2) or the simplified Eq. (4) where the phase related to the acoustic emission propagation is $\psi = 2\pi l_n / \lambda$. The path length l_n will vary only about 0.2-1.0 mm from one spatial location to the next, which is a small fraction of the spatial wavelength. Also, the spatial variations due to differences in l_n across the image may cause the phase to wrap and the variations will have a spatial wavelength close to λ .

Appendix A examines two examples of one-dimensional (1D) cases for filtering these low spatial frequency fluctuations. These 1D examples are presented to explain the concept. This concept can be readily generalized to two-dimensions, which can then be applied to images. The first model assumes that the background signal source is stationary and has constant amplitude. With a proper choice of the low-pass filter size, this background signal could be isolated. The second model assumes that the background signal source moves with the excitation source, but shifts the spatial frequency of the signal by a factor a . This factor could be related to the location of the background source such as the parametric mixing which may occur along the path to the focal point but not at the focus. Again, with appropriate filter choice, the background component can be isolated from the desired signal. These 1D models demonstrate that low-pass filtering can be used to isolate the background signals from the desired signals.

Therefore, a method that could locally estimate the background signal in the two-dimensional images and remove that background component could be used for enhancing the contrast in VA images. The acoustic emission is a sinusoidal signal, and the magnitude images only use one piece of information about the signal, because the phase is neglected. In this article, we will use the phase information in concert with the magnitude information to characterize the background, and eliminate the background signal.

2.3.2 Complex Background Suppression—We have defined the magnitude at each location above as $M(x,y)$, and the phase as $\phi(x,y)$ we can express the complex image as

$$C(x,y) = M(x,y) e^{-i\phi(x,y)}. \quad (7)$$

We need to estimate the background to identify what is associated with just the background and what relates to the object being scanned. The background in VA images is considered to be relatively constant or slowly varying because of the few number of spatial wavelengths within an image, and therefore would contain only very low spatial frequencies. One way to extract this information, is to perform spatial low-pass filtering with a convolution of the image with a large uniform kernel

$$C_f(x,y) = C(x,y) \otimes L(x,y), \quad (8)$$

$$L(x,y) = \frac{1}{W^2} J_W, \quad (9)$$

where J_W is a square matrix of ones of dimensions $W \times W$.

If the kernel size is small, then the background estimate will be very localized and the background estimate will be very close to the original image. Thus, a significant portion of the original will be identified as background and subtracted. At the other extreme, if the kernel is close to the full size of the image, only a very coarse estimate of the background can be made, and little subtraction and contrast enhancement will result. Therefore, to characterize the slowly varying background, it is necessary to use a window that is relatively large to estimate the background used for subtraction. This will insure that $C_f(x,y)$ is a good estimate of the local background for each pixel, and subtraction of this image will yield a contrast-enhanced image.

To guide this choice of window size, we can use ratio of the object sizes in the image to the wavelength of the acoustic waves at f which assuming $c = 1540$ m/s and $f = 50$ kHz, $\lambda = 30.8$ mm. Assuming that the pixels are spaced such that $\Delta x = \Delta y = 0.2$ mm, then one wavelength corresponds to 154 pixels. If the size of the features in the images is converted into wavelengths of the acoustic emission, then the choice of W can be better assessed particularly in light of the previous theoretical development that relates the phase changes as a fraction of the acoustic wavelength. The value of W should be larger than the size of the features within an image so that a portion of the feature is not considered to be background. These feature sizes may vary based on the VA application. Considering that the variations of the background are on the order of λ , $W = \lambda$ can serve as a practical upper limit. The window size applied in the low-pass filtering step was varied in its size to explore the influence of the effects of this window size in a later section.

The result from the low-pass filter step is subtracted from the original image and the corrected magnitude and phase images can be formed from the complex data.

$$I(x, y) = C(x, y) - C_f(x, y) \quad (10)$$

$$M_I(x, y) = \sqrt{\text{Re}(I(x, y))^2 + \text{Im}(I(x, y))^2} \quad (11)$$

$$\phi_I(x, y) = \tan^{-1} \left(\frac{\text{Im}(I(x, y))}{\text{Re}(I(x, y))} \right) \quad (12)$$

This method can be considered shift invariant. As a result, the subtraction of the low-pass filter result can be thought of as applying a high-pass filter operation where

$$I(x, y) = C(x, y) \otimes (1 - L(x, y)) = C(x, y) \otimes H(x, y), \quad (13)$$

where $H(x, y) = (1 - L(x, y))$ is the high-pass filter.

We can interpret this process by examining the data in the complex plane where the real value is on the horizontal axis, and the imaginary value is on the vertical axis as shown in Fig. 3. Figure 3(a) shows the result of Eq. (2), where the signal source is the vector sum of the various VA signal sources for a single position. In this case the background signal S_B is $S_B = S_{PM} + S_{NS}$. For an image, we would have a collection of vector sums with complex values. These complex values can be represented as a cluster of points centered about some nonzero combination of real and imaginary values as depicted in Fig. 3(b). For simplicity, let us assume that S_B is constant for all points in the image. In the background estimation process, we are attempting to estimate the nonzero value, S_B , represented by the arrow in Fig. 3(b). The complex background suppression (CBS) serves to recenter the complex signal space to the origin as shown in Fig. 3(c).

An additional metric to use for the evaluation of the CBS algorithm with respect to the choice of W is the spread of the distribution of the complex values. We use the standard deviation of the magnitude values, $M_I(x, y)$ for each value of W and use this value, σ_W , to compare with the image results. The σ_W is related to the resulting contrast of the images after the CBS algorithm is applied. If a small value of W is used, then most of the pixels in the window are considered background pixels and the resulting image after CBS will have very low contrast and low value of σ_W . Using a large value of W may result in a poor estimation of the local background and so less suppression of the background may occur and the value of σ_W will typically be large. It is of interest to find an optimal value of W that generates improved contrast, and σ_W is calculated and evaluated as a metric for finding this optimal value of W .

2.3.3 Magnitude Background Suppression—It is important to appreciate that the CBS process described above utilizes both the magnitude and phase information. To gain a better insight into how the CBS works, we can also use the magnitude data only and perform the background suppression after applying the low-pass filter to the magnitude image only. Images formed using this magnitude background suppression (MBS) can be used for

comparison purposes to elucidate the differences between working with only the magnitude data and the full complex data.

$$M_f(x, y) = M(x, y) \otimes L(x, y) \quad (14)$$

$$M_M(x, y) = M(x, y) - M_f(x, y) \quad (15)$$

3. Experiments

We describe three experiments that were performed to evaluate the CBS algorithm using VA images obtained by scanning a sphere phantom, a urethane breast phantom, and *in vivo* human breast tissue.

3.1 Sphere Phantom

A gelatin phantom containing three 440-C stainless steel spheres was made. The spheres had diameters of 1, 2, and 3 mm. The gelatin phantom was made from 300 Bloom gelatin and powder and glycerol (Sigma-Aldrich, St. Louis, MO) with both elements added at a concentration of 10% by volume. Additionally, potassium sorbate (Sigma-Aldrich, St. Louis, MO) was added with a concentration of 10 g/L to the phantom for preservative. A thick layer of the gelatin was poured and cured. Then, a thin layer of the liquid gelatin was then poured, and the spheres were placed within the thin layer. This thin layer melts the top portion of the gelatin and bonds the two layers and holds the spheres in place. The remaining gelatin was poured over the spheres to complete the phantom. The gelatin phantom was stored in a refrigerator.

The phantom was scanned with a two-element, 3 MHz confocal transducer with a spherical focus of 70 mm. This transducer was used for all of the experiments. Each element was driven with a different signal, $f_1 = 3.11$ MHz and $f_2 = 3.05$ MHz so that $f = 60$ kHz. The toneburst length was 100 μ s. The transducer and the phantom were mounted in a large water tank ($100 \times 64 \times 36$ cm) filled with degassed water. The transducer was raster scanned over the phantom with a pixel spacing of 0.2 mm in a C-scan format. The transducer was focused at several planes in the axial direction. The phantom was scanned from $z = 45$ mm to 95 mm in 5 mm increments, where the plane of the spheres is define as $z = 70$ mm. An illustration of the phantom is shown in Fig. 4. The MBS and CBS algorithms were applied to the acquired image data with a window size of $W = 51$ ($W = 0.40\lambda$). The diameter of the largest sphere was about 15 pixels, so the value of W was chosen so that it was much larger than this image feature.

To understand how the VA image results using the original, MBS, and CBS images correlate to the PSF of the confocal configuration, we used simulation results from Fig. 1 along the central axis of the confocal transducer. This simulation result was compared to the data from the 1 mm diameter sphere as it approximated the PSF at each of the depths.

To explore the effects that this algorithm has on the contrast of the images, we examined profiles through each of the spheres in the focal plane of the transducer. The contrast-to-

background ratio (CBR) for spheres in the original, MBS, and CBS images were computed where the CBR is a modification of the contrast-to-noise ratio [30] and is written as

$$CBR = \frac{|\mu_s - \mu_b|}{\sigma_b}, \quad (23)$$

and the parameters μ_s and μ_b are the mean values of the sphere and background and σ_b is the standard deviation of the background. We can also calculate the absolute contrast (AC) as $|\mu_s - \mu_b|$. The mean for the sphere was computed within the -6 dB limits of the profile, and the background mean and standard deviation was computed from an area located in the gelatin block and away from the sphere. The CBR was also computed when the CBS algorithm was applied with window sizes of $W = 21, 51$, and 81 ($W = 0.16\lambda, 0.40\lambda, 0.63\lambda$).

3.2 Breast Phantom

A urethane breast phantom (CIRS Model 013, Norfolk, CT) was placed in a large water tank and the confocal transducer was mechanically raster scanned across the focal plane to create an image. The ultrasound frequencies used were $f_1 = 3.10$ MHz and $f_2 = 3.05$ MHz for $f = 50$ kHz. The breast phantom was scanned at different depths with respect to the front surface of the phantom. An illustration of the phantom from two views is shown in Figure 5 drawn to scale. This diagram depicts the three-dimensional distribution of the lesions embedded within the phantom. The inclusions are also labeled for later reference and contrast evaluation. The acquired image data were processed using the CBS algorithm using a value of $W = 61$. To explore the effects that the value of W has on the data, the images at $z = 10$ and 35 mm were processed with W values of $31, 61$, and 121 ($W = 0.20\lambda, 0.40\lambda, 0.79\lambda$).

3.3 In vivo Breast Imaging

Data acquired from *in vivo* vibro-acoustic breast imaging using the apparatus described by Alizad, *et al.*, [31] was used for processing with the CBS algorithm. This device combines a VA system and a stereotactic X-ray imaging system [31]. The transducer is scanned within a small water tank and transmits ultrasound waves through a membrane into the woman's breast. The hydrophone is coupled to the breast using ultrasound gel. The transducer used in this system is identical to the one used in the sphere and breast phantom experiments. For this data, the ultrasound frequencies used were $f_2 = 3.00$ MHz and $f_1 = 3.05$ MHz in one experiment giving $f = 50$ kHz, and in the other experiment $f_1 = 3.06$ MHz so $f = 60$ kHz. The acquired image data were processed using the CBS algorithm using different values of W .

4. Results

Figure 6 shows the results of imaging the sphere phantom with the focus of the transducer at different planes. The figure shows the original images formed using the magnitude data of the complex data and images processed using the MBS and CBS algorithms. The images were scaled independently for each depth. The dynamic ranges of the images varied from 40 - 55 dB. The background intensity in the original images is quite high. The background levels of the MBS and CBS images are very low and appear black. When the CBS algorithm is applied, the contrast increases substantially, and the spheres cannot be seen at the very

shallow and deep imaging planes. While the MBS algorithm was able to reduce the background levels to low levels, the contrast was only enhanced maximally through application of the CBS algorithm. These results substantiate that the phase, which is ignored in the MBS algorithm, is important and should not be ignored. One notable change was the contrast inversion, compared to the original images, of the smallest sphere at 75 mm as well as the large sphere at 80 and 85 mm. This happens to a limited degree in the MBS images, but is present in these three cases in the CBS images. Using the complex data is a more complete approach and can yield superior results that can cause a contrast inversion making features that were once darker than the background now brighter than the background. The use of the phase and magnitude to characterize the background shifts the complex pixel value distributions and can result in changing the signal intensity with respect to the background.

Figure 7 shows the effects of the MBS and CBS algorithms on the complex data for the image at $z = 70$ mm. Each circle represents the real and imaginary components of the complex value for each pixel in the image. It is evident that in the complex signal space for the original image in Fig. 7(a), that the distribution of pixels is shifted away from zero to the left on the real axis. Although, the MBS algorithm centered the distribution around zero, the spread is reduced, that is, the brightness bias is removed, but contrast is not significantly enhanced. Once the CBS algorithm is applied the distribution becomes centralized about the origin, but maintains the overall spread of the original data, that is, the brightness bias is removed and the contrast is enhanced. The shift in the complex signal space is evident in the images by the transition from a gray background to one that is nearly black. This empirical result matches the theoretical model depicted in Fig. 3.

We examined profiles through each of the spheres for the original, MBS, and CBS images in Fig. 8. The MBS algorithm serves to shift the profile from the original image downward so that the background is near zero. The CBS algorithm maintains the peak's high level, but also brings the background level down to zero. As a result the contrast is significantly improved. The CBR and AC were computed for the image at $z = 70$ mm for each sphere in the original, MBS, and CBS images and tabulated in Table 1. Masks were applied for each sphere and regions above the spheres were used for calculating the characteristics of the background. The CBR is highest for the CBS images and increases with sphere diameter. The AC is lower in the MBS images than the original images, but the AC increases substantially when the CBS algorithm is applied.

The only parameter that can be varied in the CBS algorithm is the window size, W . We varied the value of W to 21, 51, and 81. The resulting profiles for the different spheres in the image at $z = 70$ mm are plotted in Fig. 9 as well as the values of σ_W for a range of $W = 11-101$. The profiles for $W = 51$ are somewhat difficult to distinguish from those when $W = 81$ was used. For the 3 and 2 mm diameter spheres, the larger windows improved the contrast. For the 1 mm diameter sphere, the W values used did not change the results significantly. The CBR and AC values for the different spheres and different values of W are summarized in Table 2. The CBR values improve with increasing window size and increasing sphere diameter. The maximum AC for each sphere changes with window size. The AC is maximum for the 1, 2, and 3 mm diameter spheres with W values of 21, 51, and

81, respectively. This result shows that features with different sizes have enhanced contrast with different windows. It should be noted, that the changes in most cases are not significantly large. The σ_W values rise and then plateau indicating that the spread of the complex values is very low with small W because, almost all pixels are considered background and is filtered, but as W increases the features in the image are preserved and the background is better identified and the intensity values are decreased to a near constant value. For the smaller spheres, there is less variation in the σ_W values so this metric is not as effective.

We used the mean values of the signal intensity from the 1 mm diameter sphere in each of the imaging planes for the original, MBS and CBS images to obtain an approximation of the excitation field produced by the VA method, similar to examining the PSF along the axis of the transducer. The ideal response was numerically evaluated (Fig. 1(c) at $x = 0$ mm), and the simulated result is shown with the experimental results in Fig. 10. All results were normalized independently for each set of images. Only the CBS result agrees with the simulation result. The original and MBS results show a very low result for $z = 70$ and 75 mm imaging planes. This result serves to validate that the CBS algorithm provides results that better resemble the PSF of the imaging system.

The profiles of the spheres also provide insight into how the MBS and CBS algorithms act on the data to improve contrast, but the CBS method obtains both a near zero level in the background while retaining the peak of the data found in the original data. The CBR and AC values in Table 1 show quantitatively that the CBS results maximize the contrast. The absolute contrast was interesting in that the MBS algorithm produced higher AC values than in the original images, but the CBS images had substantially higher AC values. There was also an observed trend that the larger the target, the better the contrast became. Varying the window size produced different values of the CBR, and a trend was observed with increased contrast from larger window sizes and larger targets. The AC was not as sensitive to window size, but the AC was optimized for different size spheres using different window sizes. In the optimal case for the AC, the ratio of the window size to the sphere diameter was 4.20, 5.10, and 5.40 for the spheres of 1, 2, and 3 mm diameter.

Figure 11 shows imaging results from the breast phantom at different planes in the breast phantom. The imaging planes were referenced to the front surface of the phantom and range from 10-35 mm from the surface. There is a substantial background signal from the phantom even though the background material is urethane rubber with no ultrasonic scatterers. In different images, the edges of the lesions become sharp based on the depth of the lesion and the imaging plane. Figure 12 shows the same images after applying the CBS algorithm. The contrast significantly improves and the lesions that are brightest in each image also have the sharpest edges, indicating that they are in focus. The progression of the lesions coming in and out of focus correlates with the depth diagram of the images in Fig. 5(c). Additionally, because of the high background levels in the original images, there are some contrast inversions in the CBS images when compared with the original images. The images in Figs. 11 and 12 were independently scaled for each subpanel and the dynamic range for the images was 40-48 dB.

We calculated the CBR and AC values from the data in Figs. 11 and 12 for each inclusion as labeled in Fig. 5(c) and the values were summarized in Tables 3 and 4, respectively. A mask was created for each inclusion and two rectangular background regions that were situated between inclusions 5 and 6 and 6 and 7 and the pixel values in those masked and background regions were used to calculate the means and standard deviation used in the CBR and AC equations.

We also explored how window size changes the appearance of the images processed with the CBS algorithm, and those results are shown in Fig. 13. We processed the images at $z = 10$ and 30 mm from the phantom surface with windows of size $W = 31, 61$, and 121 . As the window size increases, the lesions that are in focus tend to have a more uniform gray level, whereas at $W = 31$, the interior of the lesions in focus is darker than the outside borders. Figure 14 shows the values of σ_W with varying values of W for the breast phantom images at $z = 10$ and 30 mm. The values increase and have a small plateau and then continue to increase. This increase of σ_W is reflected in the images as with an increase of the background brightness level. For an image of something similar to this breast phantom, it may be advisable to use a value of W where the rate of change of σ_W is minimal, in this case, when $W = 61$. We compiled the CBR and AC values for each inclusion in the images in Fig. 13 in Tables 5 and 6, respectively.

The breast phantom results build upon the results from the sphere phantom, providing a perspective on the performance of the CBS method with a more complicated object. The breast phantom has multiple simulated lesions for training image-guided needle biopsy. Because the lesions are at various depths, the lesions appear more sharp and brighter when in focus and darker and less sharp when outside of the focal plane of the transducer. This can be appreciated to a certain degree in the images in Fig. 11, but better appreciated after CBS processing as shown in Fig. 12. Lesions become brighter as they come into the focal plane of the transducer. This pattern can be compared with the diagram in Fig. 5 which depicts the three-dimensional distribution of the simulated lesions. Another comparison of window size variation was shown in Fig. 13. As the window size increases, the interior of the lesions becomes more uniform in intensity, but the overall background also increases, which can reduce overall image contrast.

We applied the CBS algorithm to two different data sets from *in vivo* breast scanning. Both patients had fibroadenomas. In Fig. 15, the original image generated at $f = 50$ kHz shows the fibroadenoma as a dark region in the upper central part of the image. This image was processed with the CBS algorithm a window size of $W = 41$ ($W = 0.27\lambda$) and the results are shown in Fig. 15. In the CBS images, the fibroadenoma becomes brighter than the background, and the background becomes dark. We evaluated the CBR and AC values for the images in Fig. 15. A mask was used for the lesion and a rectangular background region to the right of the lesion was used for the background characterization. The CBR values for the original and CBS images were 1.02 and 1.49, respectively. The AC values for the original and CBS images were 24.51 and 16.07, respectively.

Figure 16 shows the results from CBS processing of an image with a large calcified fibroadenoma that was imaged using $f = 60$ kHz. The window sizes used for this image

were $W = 61, 101$, and 141 ($W = 0.48\lambda, 0.79\lambda, 1.10\lambda$). The contrast improves with the CBS processing and increasing the window size serves to make the interior of the lesion more uniform. Table 7 summarizes the CBR and AC values for the images in Fig. 16. A mask was used for the lesion and a rectangular background region to the right of the lesion was used for the background characterization. Figure 17 shows the values of σ_W with varying values of W for the image of the calcified fibroadenoma in Fig. 16. Similar to the plots for the breast phantom in Fig. 14, the values increase and have a small plateau and then continue to increase. This increase of σ_W is reflected in the images as with an increase of the background brightness level. Each of the images in Figs. 15 and 16 were independently scaled, and the dynamic range for the images was 40-48 dB.

5. Discussion

We examined the sources of the acoustic pressure in a VA setting with frequency f from a conceptual perspective. We used this analysis to understand the source(s) of the nonzero background signal found in a majority of VA images. We described experiments with a gelatin phantom with spherical inclusions as well as a commercial urethane breast phantom with irregularly shaped inclusions. In the case of these phantoms, where the background material is uniform and does not contain any scatterers, the parametric mixing signal source would seem to be dominant because the radiation force and nonlinear scattering are strongest in the presence of an object that will strongly interact with the ultrasound waves. In the case of the parametric mixing component from pixel-to-pixel, the magnitude and phase of that signal should be very similar. The only variations in the magnitude and phase may result from the variation in distance that the signal must travel from the excitation point to the hydrophone. Considering the long wavelength of the acoustic signal at frequency f , such variations should occur at low spatial frequency. This is the main reason we can separate the background from the image features in the object. These effects would be captured by the spatial filtering step in the complex background suppression.

In the case of tissue, background signal could come from other mechanisms such as the radiation force and nonlinear scattering, and there will be more variations in the magnitude and phase. The background will not reduce to zero, but features that are in focus will be enhanced more preferentially.

One of the most important aspects of the CBS algorithm is that it produces VA images with significantly enhanced contrast. The collective imaging results from two different phantoms and *in vivo* VA images demonstrate that the contrast can be improved and that objects or features that are located in the focal plane of the transducer are made brighter than other objects that may not be in the focal plane. The improvement in contrast and the ability to better localize different objects in VA images serves to improve the overall image quality. In some cases we observed a contrast inversion, that is, a feature that was darker than the background became brighter than the background in the processed images.

The sphere phantom was an idealized object which provided a good tool to learn how the CBS algorithm improved images. Both the MBS and CBS algorithms improved the contrast of the VA images, both in a qualitative and quantitative sense. The MBS algorithm

improved the contrast primarily by reducing the background signal to a near zero level. However, the CBS algorithm provided even better contrast compared to the MBS algorithm.

These changes in the images can be better understood from the complex plane analysis in Fig. 7. The background subtraction centers the distribution of the pixel values, but only with the CBS algorithm is the full dynamic range of the data retained. The MBS and CBS algorithms both center the data, but the phase information is required in the CBS algorithm to obtain maximal contrast or spread of the pixel distribution.

We compared the responses at different imaging planes with the theoretical PSF response for VA performed with this confocal transducer. In the original and MBS images in Fig. 6, at $z = 75$ mm, the response from the 1 mm diameter sphere is lower than the background signal. Looking at the corresponding CBS image at the same depth, the sphere appears brighter than the background, which would be expected. This contrast inversion with respect to the background is an artifact, which is caused by the background signal being almost 180° out-of-phase with respect to the signal from the sphere. This artifact is corrected by the CBS processing. The results from the CBS method are validated in the comparison between the responses of the 1 mm diameter sphere at different imaging planes compared to the theoretical PSF function along the axis of the transducer. The CBS algorithm, makes the response always positive, and matches very well with the theoretical curve. Another aspect is that the overall image intensity of an object will increase to its maximal value when it is in focus. This allows for better localization of an object within a volume, which can be a complicated task as illustrated by a previous study by Silva, *et al* [23]. These results imply that the complex signal is integral for image formation in VA and for localizing the object.

In the breast phantom images, we found that the signal intensity and contrast from each inclusion was maximized when the inclusion was in the focal plane of the transducer. The images in Fig. 12 and the contrast results summarized in Tables 3 and 4 demonstrate this characteristic qualitatively and quantitatively. The edges of the inclusions are maximally sharp when the inclusion is in the focal plane as well.

In the sphere and breast phantoms, there ring-like features were observed around the spheres and inclusions. These features are present in the original images, but are more emphasized by the CBS algorithm. The ring feature is a known artifact associated with the confocal transducer performing VA imaging [10, 32]. The point-spread function has been shown to act similar to a differentiator or “edge enhancer”, so the edge is defined with a halo or ring around a lesion or feature by the point-spread function and not the filter employed by the CBS or MBS algorithms.

Lastly, we used the CBS method on images of breast fibroadenomas from *in vivo* VA scanning from two different patients. In the example in Fig. 14, the fibroadenoma appeared darker than the surrounding tissue. After applying the CBS algorithm, the lesion appeared brighter than the background, but that is probably due to the fact that the CBS algorithm serves to make only what is in focus brighter. In a sense, the method serves as a deconvolution filter to provide better depth localization for the image observer. When the window size increased, more of the surrounding tissue seemed to increase in intensity.

In the case in Fig. 15, the fibroadenoma was as bright as the surrounding tissue. The CBS algorithm provided enhanced contrast both around the lesion and in the entire image so that underlying tissue structures could be appreciated more distinctly. When the window size was small, the lesion intensity value was not uniform, but that improved with increased window size. The CBS algorithm also corrected the non-uniform image intensity in the lower right corner and improved contrast in this corner.

Overall, the CBS algorithm provides better image quality for VA images, improving both the contrast and ability to spatially locate objects. The only variable parameter is the window size used for the low-pass filtering step. We have demonstrated the effects of window size in the 1D examples in Appendix A and experimentally in various phantoms and found that different images may require different window sizes for optimal image enhancement. The optimal size is dependent on the relative sizes of the image features in each image, and could be analyzed with respect to the ratio of feature size to the wavelength of the f waves. The window sizes ranged from $0.16-1.10\lambda$ in this study. An optimization scheme is a subject of further study for broader implementation of this method for VA imaging. A statistical approach for characterizing the background could be used to determine an optimal window size or an optimal window size that varies spatially. Additionally, an analytical formulation in two- or three-dimensions may aid in optimizing the application of the CBS algorithm for suppressing the influence of the background signals incorporating aspects of Eqs. (2)-(4) and the example given in Appendix A.

The algorithm does not take much more added time for processing than conventional magnitude VA image processing, and requires only simple linear signal processing software to produce the results.

Though there are multiple sources that may contribute to the background signal, the CBS algorithm does not depend on trying to estimate the different sources. The background signals all occur at the frequency of interest, f , so separating the sources in the time- or frequency-domain is difficult. However, the spatial low-frequency characteristic of the background allows the CBS algorithm to be optimally applied. CBS performance depends on the characteristics of the object. The algorithm works best when there is a relatively large and uniform background available for CBS to have an accurate estimate of the background. This often happens when imaging a lesion in a background tissue. However, if there are many structures within the imaging area, then the CBS algorithm may not perform as optimally. Therefore, performance of the CBS algorithm is limited for non-uniform objects. Also, because the out-of-focus structures are normally faded, they blend in the background. Hence, for practical purposes, CBS correction is limited to the structures that are in the focal plane.

Another approach to reducing the background in VA images is to avoid the interaction of the two beams before the desired focal location, which would require large distances or angles to be used to separate the two sources for f_1 and f_2 . Reliable ways to perform this task can be realized by using two separate transducers in an X-focal configuration [26] or by creating beams separated by large distances using a linear array [25, 29]. However, the X-focal configuration is difficult to use because alignment of the two sources is tedious. Using large

separations with the linear array system can reduce the number of elements used to produce the apertures for f_1 and f_2 and decrease the interaction region of the ultrasound beams, and thereby reduce the signal-to-noise ratio in the resulting image.

Another complication that can be encountered in VA experiments is acoustic reverberation of the aforementioned signals that can produce time-delayed replicas of the primary signals. When experiments are performed in a large water tank (typical size: $100 \times 64 \times 36$ cm), reflected acoustic waves must travel long distances before being received at the hydrophone. This time delays which range from 240-667 μ s are typically enough to separate direct propagation of the primary source acoustic emission from the focal point of the transducer to the hydrophone. As a result the contributions from that may occur are gated out because they happen later in the time recording. This particular simplification may not be valid in the *in vivo* scanning situation. For example, in VA breast imaging, the acoustic field produced at a spatial location may propagate to the skin surface where there is a tissue-air interface and a reflection of the acoustic signal is likely to occur. Within the confines of the human breast, this reverberation time will be much shorter than that experienced in the large water tank. This reverberation was studied in some detail by Urban, et al., in a recent study [24]. Further study will be devoted to applying the CBS algorithm to data that includes acoustic reverberation.

This general CBS algorithm could also be applied outside the realm of VA imaging into areas of physics and engineering that have narrowband signals at a certain frequency f , but there may be a direct signal as well as extraneous interfering signals that occur at the same time. Areas to consider would be acoustic and electromagnetic waves in sonar or radar, respectively. The CBS algorithm could be used to characterize the background interfering signals and extract and suppress their influence for measurement or imaging.

6. Conclusion

We have presented a new algorithm which uses the complex data from the acoustic emission measured in vibro-acoustography imaging. The algorithm suppresses the background contribution of the images producing images with enhanced contrast and better ability to distinguish object depth. The algorithm was tested on VA imaging results from a sphere and breast phantom as well as *in vivo* breast images. In all cases, the CBS algorithm produced images with improved quality. This method is a fast, linear approach towards improving the images produced with vibro-acoustography.

Acknowledgments

This work was supported in part by grants R01CA127235, P50CA91956, R21CA121579, R01CA148994, and R01CA168575 from the National Institutes of Health and grant BCRT0504550 from the Susan G. Komen Breast Cancer Foundation. The authors thank Dr. James F. Greenleaf for insightful discussions, Randall R. Kinnick and Thomas M. Kinter for laboratory support, and Jennifer Milliken for administrative support. MWU is grateful to Dr. Farid Mitri for helpful correspondence on vibro-acoustography signal sources. Disclosure: Mayo Clinic and two of the authors (MWU and MF) have a financial interest associated with technology used in this research; the technology has been licensed in part to industry.

Appendix A

This one-dimensional illustration describes how the CBS method works to filter a desired signal combined with an undesired background signal. In this illustration we will ignore the temporal component of the signal and assume that all signals that we are interested in are at frequency ω . We define the spatial variation of the pressure, or signal, received at the hydrophone as

$$s(x_0) = s_0(x_0) e^{jk_0 x_0} + s_b(x_0) e^{jk_0 x_b}. \quad (\text{A.1})$$

where $s_0(x_0)$ is the spatial amplitude distribution for the true signal that we want to extract and enhance, $s_b(x_0)$ is the spatial amplitude distribution for the background signal that we want to eliminate. The spatial frequency corresponding to ω is $k_0 = \omega/c$. The path lengths for the desired and background signals are x_0 and x_b , respectively. We will assume that x_b is fixed and $s_b(x_0)$ is constant. This implies that the background signal is a fixed source with respect to the hydrophone.

$$s(x_0) = s_0(x_0) e^{jk_0 x_0} + s_b e^{jk_0 x_b} \quad (\text{A.2})$$

The second term is constant and has a spatial frequency of zero. As a result, the magnitude of the signal will vary with x_0 . If we assume that s_0 is constant then the magnitude of $s(x_0)$ will oscillate between a maximum of $|s_0 + s_b|$ and a minimum of $|s_0 - s_b|$. The spatial distance between the peaks is $k_0 x_0 = 2\pi$ and solving for $x_0 = 2\pi/(2\pi f/c) = c/f = \lambda$. This variation represents the background artifactual variation.

If the path length for the background signal varies as a function of x_0 , for example if $x_b = ax_0$, and assuming that $a \neq 1$. If $a = 1$, then the magnitude of $s(x_0)$ will not have a cyclic variation with x_0 . If $a < 1$ the spatial cyclic variations will be larger than λ .

$$s(x_0) = s_0(x_0) e^{jk_0 x_0} + s_b e^{jk_0 a x_0} \quad (\text{A.3})$$

The spatial frequency of the desired signal is k_0 while the spatial frequency of the background signal is now ak_0 and is lower than the spatial frequency of the desired signal.

Our aim is to eliminate this background signal. One approach is to take a spatial average of the image or 1D signal in this illustration and subtract it out, or equivalently, extract the high-pass portion of the spatial frequency domain.

We transform the measured signal from the spatial domain to the spatial frequency domain where the Fourier transform of $s(x_0)$ is $S(k)$ and

$$S(k) = S_0(k - k_0) + S_b \delta(k - ak_0). \quad (\text{A.4})$$

We can see that second term is an impulse function that is positioned at $k = ak_0$, and the desired signal is centered at k_0 . If we apply a low-pass filter by convolving $s(x_0)$ with a rect function of width, W we obtain

$$s_{LP}(x_0) = [s_0(x_0)e^{jk_0x_0} + s_b e^{jk_0ax_0}] \otimes \text{rect}\left(\frac{x}{W}\right). \quad (\text{A.5})$$

This relationship in the spatial frequency domain can be written as [33]

$$S_{LP}(k) = [S_0(k - k_0) + S_b \delta(k - ak_0)] \cdot W \text{ sinc}(Wk). \quad (\text{A.6})$$

The main lobe of the sinc function acts as the low-pass filter so if the width of this lobe is wide enough to encompass the background signal, this signal will be extracted from the measured signal.

Examples of applying filters of different sizes are shown in Fig. A.1. In Fig A.1(a) the model from Eq. (A.2) where $s_0 = 2$, $s_b = 0.6$, and $x_b = 100$ mm is shown with filters of length $W = 81$ and 141 . This case could also be generalized for Eq. (A.6) by setting $a = 0$. There is a background signal component at $k = 0$ mm⁻¹ and the desired signal component at $k = 34$ mm⁻¹. In Fig A.1(b) the model from Eq. (A.3) which is realized in the spatial frequency domain in Eq. (A.6). For this case, $s_0 = 2$, $s_b = 0.6$, and $a = 0.6$ is shown with filters of length $W = 81$ and 141 . There is a background signal component at $k = 16$ mm⁻¹ and the desired signal component at $k = 34$ mm⁻¹. In both models, the filter with $W = 81$ will diminish the desired component as well as filter out the background signal, while the filter with $W = 141$ will spare the desired signal component and only filter out the background signal.

References

1. Fatemi M, Greenleaf JF. Ultrasound-stimulated vibro-acoustic spectrography. Science. 1998; 280:82–85. [PubMed: 9525861]
2. Fatemi M, Greenleaf JF. Vibro-acoustography: An imaging modality based on ultrasound-stimulated acoustic emission. Proc. Natl. Acad. Sci. U. S. A. 1999; 96:6603–6608. [PubMed: 10359758]
3. Pislaru C, Kantor B, Kinnick RR, Anderson JL, Aubry MC, Urban MW, Fatemi M, Greenleaf JF. In vivo vibroacoustography of large peripheral arteries. Invest. Radiol. 2008; 43:243–252. [PubMed: 18340248]
4. Alizad A, Fatemi M, Wold LE, Greenleaf JF. Performance of vibro-acoustography in detecting microcalcifications in excised human breast tissue: a study of 74 tissue samples. IEEE Trans. Med. Imaging. 2004; 23:307–312. [PubMed: 15027523]
5. Alizad A, Whaley DH, Urban MW, Carter RE, Kinnick RR, Greenleaf JF, Fatemi M. Breast vibro-acoustography: initial results show promise. Breast Cancer Res. 2012; 14:R128. [PubMed: 23021305]
6. Alizad A, Urban M, Morris J, Reading C, Kinnick R, Greenleaf J, Fatemi M. In vivo thyroid vibro-acoustography: a pilot study. BMC Medical Imaging. 2013; 13:12. [PubMed: 23530993]
7. Mitri FG, Davis BJ, Alizad A, Greenleaf JF, Wilson TM, Mynderse LA, Fatemi M. Prostate cryotherapy monitoring using vibroacoustography: preliminary results of an ex vivo study and technical feasibility. IEEE Trans. Biomed. Eng. 2008; 55:2584–2592. [PubMed: 18990628]

8. Mitri FG, Davis BJ, Urban MW, Alizad A, Greenleaf JF, Lischer GH, Wilson TM, Fatemi M. Vibro-acoustography imaging of permanent prostate brachytherapy seeds in an excised human prostate -Preliminary results and technical feasibility. *Ultrasonics*. 2009; 49:389–394. [PubMed: 19062061]
9. Alizad A, Mehrmohammadi M, Mitri FG, Davis BJ, Sebo TJ, Mynderse LA, Kinnick RR, Greenleaf JF, Fatemi M. Application of vibro-acoustography in prostate tissue imaging. *Med. Phys.* 2013; 40:022902. [PubMed: 23387773]
10. Alizad A, Wold LE, Greenleaf JF, Fatemi M. Imaging mass lesions by vibro acoustography: modeling and experiments. *IEEE Trans. Med. Imaging*. 2004; 23:1087–1093. [PubMed: 15377117]
11. Mitri FG, Fatemi M. Improved vibroacoustography imaging for nondestructive inspection of materials. *J. Appl. Phys.* 2005; 98:114901.
12. Mitri FG, Eberlein U, Fatemi M. Surface roughness imaging using the acoustic emission induced by the dynamic radiation force of ultrasound. *Appl. Phys. Lett.* 2006; 88:234105.
13. Urban MW, Alizad A, Aquino W, Greenleaf JF, Fatemi M. A review of vibro acoustography and its applications in medicine. *Curr. Med. Imaging Rev.* 2011; 7:350–359. [PubMed: 22423235]
14. Fatemi M, Greenleaf JF. Probing the dynamics of tissue at low frequencies with the radiation force of ultrasound. *Phys. Med. Biol.* 2000; 45:1449–1464. [PubMed: 10870703]
15. Silva GT, Mitri FG. Difference-frequency generation in vibro-acoustography. *Phys. Med. Biol.* 2011; 56:5985–5993. [PubMed: 21860078]
16. Thuras AL, Jenkins RT, O'Neil HT. Extraneous frequencies generated in air carrying intense sound waves. *J. Acoust. Soc. Am.* 1935; 6:173–180.
17. Westervelt PJ. Scattering of sound by sound. *J. Acoust. Soc. Am.* 1957; 29:199–203.
18. Silva GT, Chen S, Viana LP. Parametric amplification of the dynamic radiation force of acoustic waves in fluids. *Phys. Rev. Letters*. 2006; 96
19. Thierman, JS. Sources of difference frequency sound in a dual-frequency imaging system with implications in monitoring thermal surgery. Massachusetts Institute of Technology; 2004.
20. Barriere, C. Physique Macroscopique. University of Paris; Paris, France: 2001. Effet de la diffraction sur l'interaction paramétrique d'ondes acoustiques. Application à la mesure de paramètres de non linéarité et de champs acoustiques.
21. Jensen JA, Svendsen NB. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*. 1992; 39:262–267. [PubMed: 18263145]
22. Jensen JA. Field: A program for simulating ultrasound systems, in: 10th Nordic-Baltic Conference on Biomedical Imaging. Medical & Biological Engineering & Computing. 1996:351–353. [PubMed: 8945858]
23. Silva GT, Frery AC, Fatemi M. Image formation in vibro-acoustography with depth-of-field effects. *Comput. Med. Imaging Graph.* 2006; 30:321–327. [PubMed: 16949793]
24. Urban MW, Alizad A, Fatemi M. Vibro-acoustography and multifrequency image compounding. *Ultrasonics*. 2011; 51:689–696. [PubMed: 21377181]
25. Silva GT, Greenleaf JF, Fatemi M. Linear arrays for vibro-acoustography: a numerical simulation study. *Ultrason. Imaging*. 2004; 26:1–17. [PubMed: 15134390]
26. Chen S, Fatemi M, Kinnick R, Greenleaf JF. Comparison of stress field forming methods for vibro-acoustography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*. 2004; 51:313–321. [PubMed: 15128218]
27. Silva GT, Chen S, Frery AC, Greenleaf JF, Fatemi M. Stress field forming of sector array transducers for vibro-acoustography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*. 2005; 52:1943–1951. [PubMed: 16422406]
28. Urban MW, Silva GT, Fatemi M, Greenleaf JF. Multifrequency vibro-acoustography. *IEEE Trans. Med. Imaging*. 2006; 25:1284–1295. [PubMed: 17024832]
29. Urban MW, Chalek C, Kinnick RR, Kinter TM, Haider B, Greenleaf JF, Thomenius K, Fatemi M. Implementation of vibro-acoustography on a clinical ultrasound system. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*. 2011; 58:1169–1181.

30. Hooi FM, Thomenius KE, Fisher R, Carson PL. Hybrid beamforming and steering with reconfigurable arrays. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control.* 2010; 57:1311–1319. [PubMed: 20529707]
31. Alizad A, Whaley DH, Greenleaf JF, Fatemi M. Critical issues in breast imaging by vibro-acoustography. *Ultrasonics.* 2006; 44(Suppl 1):e217–220. [PubMed: 16843513]
32. Fatemi M, Wold LE, Alizad A, Greenleaf JF. Vibro-acoustic tissue mammography. *IEEE Trans. Med. Imaging.* 2002; 21:1–8. [PubMed: 11838661]
33. Bracewell, RN. *The Fourier Transform and Its Applications.* Third ed.. McGraw Hill; Boston, MA: 2000.

Highlights

- Complex Background Suppression (CBS) uses the complex acoustic emission data.
- CBS suppresses the background contribution of the images.
- Images have enhanced contrast and better ability to separate object depth.
- CBS was applied to sphere and breast phantom images and *in vivo* breast images.
- The CBS method is a fast, linear approach to improve vibro-acoustography images.

Fig. 1.

Spatial distributions of acoustic fields produced by two-element confocal transducer spherically focused at 70 mm. All images are independently normalized and have a 50 dB dynamic range, (a) Pressure field from central disc of radius 14.8 mm at $f_1 = 3.00$ MHz, (b) Pressure field from annular ring with inner radius 16.8 mm and outer radius of 22.5 mm at $f_2 = 2.95$ MHz, (c) Spatial distribution of the f component of the dynamic radiation force.

Fig. 2. Diagram of sources of VA signals. The sources are due to radiation force excitation, (s_{RF}), parametric mixing (s_{PM}), nonlinear scattering (s_{NS}), and transducer vibration (s_V). Acoustic signals are measured at the hydrophone.

Fig. 3. Complex signal domain description of CBS algorithm. (a) Vector summation of different VA signal components, (b) Pixel value locations in complex signal space with arrow denoting location of distribution mean, (c) Pixel value locations after mean signal subtraction.

Fig. 4. Sphere phantom setup. Imaging planes range from $z = 45$ - 95 mm and are separated by 5 mm. The spheres are located at $z = 70$ mm. Three stainless steel spheres were used which have diameters of 1, 2, and 3 mm.

Fig. 5.

Breast phantom imaging setup. (a) Photograph of breast phantom, (b) Diagram of scanning setup, (c) Diagram of breast phantom to scale with front and side views. The front view is scaled to be 40×80 mm and the side view is scaled to be 40×43 mm. In the side view, the left edge is the front of the phantom and is closest to the transducer. During imaging, the transducer was offset from the front edge different amounts to obtain different focal planes. Different texture patterns in the lesions are used to match lesions in side and front views. The inclusions are labeled for later reference and contrast evaluation. [Adapted with permission from [29], © IEEE 2011]

Fig. 6.

Sphere phantom image results for scanning spheres at different planes from $z = 45$ to $z = 95$ in 5 mm increments. Images have been scaled on independent scales for each depth. (a) Original images, (b) MBS images, (c) CBS images.

Fig. 7.

Complex signal space for images at $z = 70$ mm. The real component is on the horizontal axis and the imaginary component is on the vertical axis. (a) Signal space for original image, (b) signal space for image after MBS algorithm was applied, (c) signal space for image after CBS algorithm was applied.

Fig. 8. Profile plots of spheres for different algorithms from image at $z = 70$ mm. (a) 3 mm diameter sphere, (b) 2 mm diameter sphere, (c) 1 mm diameter sphere.

Fig. 9.

Image profiles and σ_W plots of spheres from image at $z = 70$ mm for different window sizes for CBS algorithm. (a) Image profiles for 3 mm diameter sphere, (b) σ_W for 3 mm diameter sphere, (c) Image profiles for 2 mm diameter sphere, (d) σ_W for 2 mm diameter sphere, (e) Image profiles for 1 mm diameter sphere, (f) σ_W for 1 mm diameter sphere.

Fig. 10.

Comparison of axial field response from Fig. 1(c) at $x = 0$ mm (solid line) with mean values of 1 mm sphere (dashed line) response at different imaging planes from $z = 45$ to $z = 95$ in 5 mm increments. (a) Original images, (b) MBS images, (c) CBS images.

Fig. 11. Original images of breast phantom at depths of 10-35 mm from surface of phantom. (a) $z = 10$ mm, (b) $z = 15$ mm, (c) $z = 20$ mm, (d) $z = 25$ mm, (e) $z = 30$ mm, (f) $z = 35$ mm.

Fig. 12.

Images of breast phantom at depths of 10-35 mm from surface of phantom after CBS algorithm was applied with $W = 61$. (a) $z = 10$ mm, (b) $z = 15$ mm, (c) $z = 20$ mm, (d) $z = 25$ mm, (e) $z = 30$ mm, (f) $z = 35$ mm.

Fig. 13.

Images of breast phantom at depths of 10 and 30 mm from surface of phantom after CBS algorithm was applied with varying window sizes. (a) $z = 10$ mm, $W = 31$, (b) $z = 30$ mm, $W = 31$, (c) $z = 10$ mm, $W = 61$, (d) $z = 30$ mm, $W = 61$, (e) $z = 10$ mm, $W = 121$, (f) $z = 30$ mm, $W = 121$.

Fig. 14. Plots of σ_W for the urethane breast phantom for images at $z = 10$ and 30 mm. The open symbols indicate the values of W used in Fig. 13.

Fig. 15.

In vivo breast image of a patient with a calcified fibroadenoma in her left breast. (a) Original image, (b) CBS image with $W = 41$.

Fig. 16.

In vivo breast image of a patient with a calcified fibroadenoma. (a) Original image, (b) CBS image with $W = 61$, (c) CBS image with $W = 101$, (d) CBS image with $W = 141$.

Fig. 17.

Plots of σ_W for the *in vivo* breast image of the calcified fibroadenoma in Fig. 16. The open symbols indicate the values of W used in Fig. 16.

Fig. A.1. Signal and filter models. (a) Model for Eq. (2) with $s_0 = 2$, $s_b = 0.6$, and $x_b = 100 \text{ mm}$ and filters with $W = 81, 141$, (b) Model for Eq. (3) with $s_0 = 2$, $s_b = 0.6$, and $a = 0.6$ and filters with $W = 81, 141$.

Table 1

Comparison of CBR and AC for different spheres and different algorithms.

Algorithm	CBR			AC		
	1 mm	2 mm	3 mm	1 mm	2 mm	3 mm
Original	0.43	4.91	7.12	12.10	139.62	202.44
MBS	5.41	22.99	26.58	32.86	139.59	202.44
CBS	24.26	48.81	61.41	149.81	301.50	379.29

Table 2

Comparison of CBR and AC for different spheres and window sizes using CBS.

Window Size	CBR			AC		
	1 mm	2 mm	3 mm	1 mm	2 mm	3 mm
21	19.27	42.25	113.74	139.83	260.93	278.25
51	19.84	48.81	156.93	143.93	301.50	383.91
81	19.71	48.78	158.68	143.04	301.28	388.20

Table 3

CBR values for original and CBS images for each depth and inclusion from image data in Figs. 11 and 12.

Inclusion	Image	z = 10 mm	z = 15 mm	z = 20 mm	z = 25 mm	z = 30 mm	z = 35 mm
1	Original	2.86	2.82	2.27	1.74	1.36	0.75
	CBS	2.45	3.08	4.03	5.01	8.16	8.93
2	Original	0.83	1.08	1.70	2.53	3.16	2.54
	CBS	17.25	8.72	2.83	0.22	0.90	0.23
3	Original	4.61	0.60	1.07	0.51	1.41	0.74
	CBS	23.79	13.68	7.49	2.07	0.57	1.33
4	Original	0.01	0.81	0.26	0.09	0.72	1.49
	CBS	5.75	4.42	5.66	6.26	9.74	10.96
5	Original	1.67	0.51	0.81	0.63	1.40	1.52
	CBS	9.30	10.93	8.48	4.63	2.41	0.94
6	Original	0.09	0.72	1.58	2.36	3.06	2.18
	CBS	4.49	8.39	11.82	14.40	16.08	11.84
7	Original	0.03	1.20	1.20	1.43	2.51	2.18
	CBS	13.87	11.57	8.49	4.42	1.85	1.14

Table 4

AC values for original and CBS images for each depth and inclusion from image data in Figs. 11 and 12.

Inclusion	Image	z = 10 mm	z = 15 mm	z = 20 mm	z = 25 mm	z = 30 mm	z = 35 mm
1	Original	41.27	57.85	47.74	42.68	23.26	16.15
	CBS	7.58	13.92	18.66	23.68	32.05	39.96
2	Original	11.98	22.15	35.82	61.98	54.10	54.49
	CBS	53.33	39.41	13.08	1.02	3.52	1.02
3	Original	66.54	12.25	22.63	12.50	24.22	15.95
	CBS	73.54	64.85	34.66	9.80	2.23	5.95
4	Original	0.18	16.72	5.47	2.14	12.25	32.09
	CBS	17.78	19.98	26.20	29.59	38.28	49.06
5	Original	24.14	10.38	17.15	15.55	24.06	32.71
	CBS	28.76	49.38	39.24	21.87	9.49	4.19
6	Original	1.31	14.84	33.24	57.95	52.41	46.86
	CBS	13.88	37.93	54.70	68.02	63.17	53.00
7	Original	0.47	24.74	25.19	34.99	43.01	46.78
	CBS	42.88	52.28	39.44	20.89	7.25	5.11

Table 5

CBR values for CBS images in Fig. 13.

Inclusion	z = 10 mm			z = 30 mm		
	W = 31	W = 61	W = 121	W = 31	W = 61	W = 121
1	5.13	2.45	1.07	14.53	8.16	2.28
2	35.79	17.25	5.65	8.83	0.90	2.07
3	46.42	23.79	9.66	5.96	0.57	1.23
4	6.59	5.75	1.15	16.08	9.74	3.86
5	16.94	9.30	1.70	6.68	2.41	0.52
6	11.98	4.49	0.50	35.01	16.08	6.90
7	24.40	13.87	3.75	6.79	1.85	1.20

Table 6

AC values for CBS images in Fig. 13.

Inclusion	z = 10 mm			z = 30 mm		
	W = 31	W = 61	W = 121	W = 31	W = 61	W = 121
1	8.17	7.58	7.16	28.08	32.05	18.87
2	56.97	53.33	37.66	11.27	3.52	17.09
3	73.90	73.54	64.37	11.53	2.23	10.15
4	10.49	17.78	7.67	31.08	38.28	31.90
5	26.97	28.76	11.30	12.91	9.49	4.32
6	19.07	13.88	3.30	67.66	63.17	57.04
7	38.84	42.88	24.96	13.12	7.25	9.91

Table 7

CBR and AC values for the original and CBS images in Fig. 16.

Image	CBR	AC
Original	3.26	39.27
W = 61	3.25	18.81
W = 101	3.82	27.29
W = 141	3.98	34.61