Published in final edited form as:

IEEE Trans Ultrason Ferroelectr Freq Control. 2017 March ; 64(3): 500–513. doi:10.1109/TUFFC.2016.2634004

Efficient Strategies for Estimating the Spatial Coherence of Backscatter

Dongwoon Hyun^{1,2}, Anna Lisa C. Crowley³, and Jeremy J. Dahl⁴

¹Department of Biomedical Engineering, Duke University, Durham, NC 27708

²Department of Bioengineering, Stanford University, Stanford, CA 94305

³Division of Cardiology, Department of Medicine, Duke University Hospital, Durham, NC 27710

⁴Department of Radiology, Stanford University, Stanford, CA, 94305

Abstract

The spatial coherence of ultrasound backscatter has been proposed to reduce clutter in medical imaging, to measure the anisotropy of the scattering source, and to improve the detection of blood flow. These techniques rely on correlation estimates that are obtained using computationally expensive strategies. In this study, we assess existing spatial coherence estimation methods and propose three computationally efficient modifications: a reduced kernel, a downsampled receive aperture, and the use of an ensemble correlation coefficient. The proposed methods are implemented in simulation and *in vivo* studies. Reducing the kernel to a single sample improved computational throughput and improved axial resolution. Downsampling the receive aperture was found to have negligible effect on estimator variance, and improved computational throughput by an order of magnitude for a downsample factor of 4. The ensemble correlation estimator demonstrated lower variance than the currently used average correlation. Combining the three methods, the throughput was improved 105-fold in simulation with a downsample factor of 4 and 20-fold *in vivo* with a downsample factor of 2.

I. Introduction

In pulse-echo ultrasound, the backscattered wave is sensed by an array of transducer elements. The signals received on the element array, or channel signals, are traditionally reconstructed into B-mode images using the classic delay-and-sum (DAS) beamformer. Time delays are applied to focus the signal at a desired direction and range, and the channel signals are summed together to measure the magnitude of the echo. DAS relies on assumptions such as a uniform speed of sound and single scattering that are often violated in medical imaging, resulting in artifacts such as phase aberrations [1] and reverberation clutter [2], [3].

Many adaptive beamforming techniques have been proposed to mitigate these effects and improve image quality. These techniques aim to extract and utilize information from the channel signals beyond the magnitude of the backscattering source. Many of these techniques rely on some measure of *spatial coherence*. Spatial coherence is an umbrella term referring to the overall similarity amongst channel signals, and is quantified by measures

such as the *spatial covariance* and the *spatial correlation* (defined in Sec. II). For example, phase aberration is often corrected by measuring the correlation between neighboring elements [4], [5]. The performance of the correction is strongly dependent on high correlation between the element signals. Metrics such as the generalized coherence factor [6] and the phase coherence factor [7] have been proposed to assess focusing quality and to reweight B-mode images. Minimum variance beamforming uses estimates of the spatial covariance matrix to suppress sidelobes and improve resolution [8]. In each of these, a form of spatial coherence is used to enhance the DAS beamformer.

Spatial coherence has been used independently of the DAS beamformer. Short-lag spatial coherence (SLSC) is a beamforming technique that reconstructs images of the spatial coherence in echoes instead of the magnitude. SLSC reduces the impact of clutter by differentiating between partially coherent tissue signals and incoherent noise regardless of magnitude [9], and has been applied successfully both in simulations and *in vivo* [10]–[13]. Spatial coherence has also been used to infer the anisotropy of the scattering source and its orientation relative to the transducer using backscatter tensor imaging (BTI) [14], and to improve the detection of blood flow by suppressing spatially incoherent noise with the coherent flow power doppler (CFPD) technique [15].

Spatial coherence imaging techniques depend on high quality estimates of the spatial coherence. For instance, the image quality in SLSC and CFPD are improved by low variance estimates, manifesting as smoother (less speckled) images with improved texture signal-to-noise ratio (SNR). In the context of spatial coherence estimation, the variance is usually improved in two ways: by using a temporal window of signal to compute the correlation and by averaging correlations with the same expected value [9], [14]–[17]. Unfortunately, these techniques are computationally expensive. The computational cost scales linearly with the size of the signal kernel used to compute the correlation and quadratically with the number of channels [17], [18]. (By contrast, DAS scales linearly with the number of channels and does not employ a kernel.) As is common of adaptive beamforming techniques, spatial coherence beam-forming currently requires either offline processing or the use of high-performance computing technologies such as GPU processing for real-time use [18]–[21]. As such, it is critical to minimize the computational impact of each coherence estimate, and to understand the fundamental sources of error in spatial coherence estimation.

In this work, we propose computationally efficient strategies to improve spatial coherence estimation. We begin with a brief assessment of the common sources of noise in coherence estimation and describe current estimation practices and identify their shortcomings and redundancies. Several efficient strategies are proposed to provide spatial coherence estimates of similar or better quality at a fraction of the computational cost of current methods, and these techniques are demonstrated in simulation and *in vivo* studies.

II. Coherence Estimation Techniques

A. Noise in Spatial Coherence Estimation

Consider two backscattered signals A and B received at two points on an aperture, modeled by zero-mean complex Gaussian random processes. Let $A[n]$ denote the n -th axial sample of

the random process A , digitized after applying geometric focal delays. The spatial covariance of A and B at sample n is given as

$$C_{AB}[n] = \langle A[n]B^*[n] \rangle - \langle A[n] \rangle \langle B^*[n] \rangle \quad (1)$$

where $*$ denotes the complex conjugate and $\langle \cdot \rangle$ denotes the expected value. The normalized covariance, referred to as the correlation, is defined as

$$\rho_{AB}[n] = \frac{C_{AB}[n]}{\sigma_A[n]\sigma_B[n]}, \quad (2)$$

where $\sigma_A[n] = \sqrt{C_{AA}[n]}$ is the standard deviation of $A[n]$. The quantities $\rho_{AB}[n]$ and $C_{AB}[n]$ are the population correlation and population covariance, respectively.

Given a finite sample of K observations, $\rho_{AB}[n]$ can be estimated using the sample correlation:

$$\hat{R}_{AB}[n] = \frac{\sum_{k=1}^K a_k[n]b_k^*[n]}{\sqrt{\sum_{k=1}^K |a_k[n]|^2 \sum_{k=1}^K |b_k[n]|^2}}, \quad (3)$$

where the k -th observation of $A[n]$ is denoted as $a_k[n]$. The process is illustrated in Fig. 1 for digitized signals. The estimates produced by (3) have some inherent variance that can be reduced by repeated observations of the signal of interest with new realizations of noise.

Noise in ultrasound imaging can be separated into two broad classes: time-*dependent* and time-*invariant* noise. Time-dependent noises ($\mathcal{T}$ noises), such as thermal noise, change on a sample-to-sample, pulse-to-pulse, or frame-to-frame basis and can be mitigated with repeated observations of the same target. With each subsequent measurement, the SNR is improved. Time-invariant noises ($\mathcal{S}$ noises) do not evolve over time and therefore cannot be reduced with repeated measurements. Examples of $\mathcal{S}$ noise include speckle, which stems from the inherent unresolvable microstructure of scatterers, and clutter, which arises from the reverberations caused by inhomogeneities in the acoustic properties of the medium. Because these are products of the physical structure of the medium, their effects can be decreased only by observing the target differently, typically by insonification from a different angle or position [22], [23] or by transmitting at a different frequency [23].

The overall variance of the spatial coherence estimates is improved only by multiple observations of both $\mathcal{S}$ and $\mathcal{T}$ noises. An infinite number of observations of $\mathcal{T}$ obtained by repeatedly imaging the target cannot eliminate the variance introduced by $\mathcal{S}$, such as the speckle pattern. Therefore, a suitable spatial coherence estimate needs to incorporate observations of both $\mathcal{S}$ and $\mathcal{T}$.

B. Estimation with a Kernel

Unfortunately, there is usually a physical limit to the number of times and ways a target can be observed. For example, physiological motion may limit a target to 2 or 3 observations, and a transducer with finite extent and bandwidth may prevent more than 3 or 4 views of the speckle target. In these instances, the correlation estimate can be improved by using a short axial window as a surrogate for more observations. The window is often referred to as a signal kernel, and is widely used to reduce estimator variance in applications including spatial coherence estimation [9], [14], [16], delay estimation [24], [25], and phase shift estimation [26], [27].

When K is limited to 1 (i.e. only a_1 and b_1 are available), a kernel of $2T + 1$ samples centered around n is implemented as

$$\hat{R}_{AB}[n] = \frac{\sum_{t=-T}^T a_1[n+t] b_1^*[n+t]}{\sqrt{\sum_{t=-T}^T |a_1[n+t]|^2 \sum_{t=-T}^T |b_1[n+t]|^2}}. \quad (4)$$

By using a signal kernel, K is artificially increased from 1 to $2T + 1$ observations. This process is illustrated in Fig. 2 for $T = 3$. It is easily shown that (4) is equivalent to the sample correlation in (3) by setting $a_k = a_1[n+k-T-1]$. In other words, neighboring samples $a_1[n+t]$ are treated as though they are additional observations of the desired signal $A[n]$, significantly increasing K . This technique is only valid when the signals $A[n]$ and $B[n]$ retain the same underlying statistics over the kernel length (i.e., $\rho_{AB}[n+t] = \rho_{AB}[n]$). In this case, each kernel sample is a new observation of $\rho_{AB}[n]$ in $\mathcal{S}$ and $\mathcal{T}$, and the kernel improves the estimate.

However, axially stationary imaging targets are not typically interesting. Most medical imaging targets have fine and intricate anatomical structures that violate this assumption, i.e., $\rho_{AB}[n+t] \neq \rho_{AB}[n]$. For such targets, the kernel samples are *not* new observations of $\rho_{AB}[n]$ in $\mathcal{S}$, but rather observations of different entities altogether. The resulting correlation estimate is an aggregation over the length of the kernel, and will manifest as a loss in axial resolution. This highlights a fundamental limitation of the kernel in spatial coherence estimation: the kernel cannot provide new observations of $\mathcal{S}$.

C. Estimation with an Average

For the received backscatter from a diffusely scattering medium, the population covariance of two aperture points is approximated by the van Cittert-Zernike theorem [16], [28]. The theorem describes the propagation of spatial covariance from an incoherent source, such as the backscatter from an insonified volume of tissue [29]. Adding a normalization term, the expected correlation (i.e. normalized covariance) of two aperture points separated by $\mathbf{r} = (x, y)$ is given as

$$\rho_{\Delta} = \mathcal{C}_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |H(u, v) \chi(u, v)|^2 \times \exp \left(-j2\pi \frac{u\Delta x + v\Delta y}{zc} \right) du dv, \quad (5)$$

where $\mathcal{C}_0$ is the expected covariance at $\Delta = 0$, (u, v) are the coordinates of the image plane (sometimes referred to as the source plane), $H(u, v)$ is the transmit beam pattern, $\chi(u, v)$ is the echogenicity pattern of the medium, z is the distance from the image plane to the aperture, and c is the speed of sound. Here, the sample number $[n]$ is omitted for clarity. For convenience, Δ is represented in units of element pitch.

The correlation in (5) depends only on the spacing of the aperture points, a property referred to as wide-sense stationarity. Consequently, the measured correlation of any two aperture points separated by Δ is an observation of ρ . Ultrasound transducers are typically composed of uniformly spaced elements, and so there are many such observations. Let ξ denote the set of all signal pairs (A, B) with the same spacing Δ :

$$\xi_{\Delta} = \{(A, B) | (x_A - x_B, y_A - y_B) = \Delta\}, \quad (6)$$

where (x_A, y_A) are the coordinates of A . For instance, a uniform linear array with 128 elements has $\xi_4 = \{(1, 5), (2, 6), \dots, (124, 128)\}$. For each pair (A, B) in ξ_4 , the measured correlation $\hat{R}_{AB}$ is a distinct (though not necessarily independent) observation of ρ_4 .

Currently, a single aggregate estimate of ρ is obtained by taking the average of the individual $\hat{R}_{AB}$ estimates [9], [14], [16], [30], [31]. The averaged correlation estimate is computed as

$$\hat{R}_{\Delta}^{\text{avg}} = \frac{1}{|\xi_{\Delta}|} \sum_{(A, B) \in \xi_{\Delta}} \hat{R}_{AB}, \quad (7)$$

where $|\xi|$ is the size of the set. Each $\hat{R}_{AB}$ is typically estimated using an axial kernel of one wavelength (1λ) [9]. Additionally, only the real component of the estimate is retained in practice. This is equivalent to combining the forward ($\hat{R}_{AB}$) and backward ($\hat{R}_{BA}$) correlations into a single correlation:

$$\text{Re} \left\{ \hat{R}_{AB} \right\} = \frac{\hat{R}_{AB} + \hat{R}_{BA}}{2}. \quad (8)$$

The variance of (7) can be written as

$$\begin{aligned}
\sigma^2 [\hat{R}_{\Delta}^{\text{avg}}] &= \frac{1}{|\xi_{\Delta}|^2} \sum_{(A,B)} \sum_{(C,D)} \text{cov} [\hat{R}_{AB}, \hat{R}_{CD}] \\
&= \frac{\sigma^2 [\hat{R}_{AB}]}{|\xi_{\Delta}|} + \frac{\sum_{(A,B)} \sum_{(C,D) \neq (A,B)} \text{cov} [\hat{R}_{AB}, \hat{R}_{CD}]}{|\xi_{\Delta}|^2}, \quad (9)
\end{aligned}$$

where each summation is over the set ξ and $\text{cov}[\hat{R}_{AB}, \hat{R}_{CD}]$ is the covariance amongst the correlations themselves. In (9), correlations are assumed to have the same variance $\sigma^2[\hat{R}_{AB}]$.

At best, the variance of $\hat{R}_{\Delta}^{\text{avg}}$ is $|\xi|$ times smaller than that of an individual estimate when each observation is uncorrelated with the others. At worst, every observation is perfectly correlated with every other observation, and there is no reduction in variance.

The averaging process is analogous to spatial compounding of B-mode images. Spatial compounding averages multiple images of the same target viewed from different spatial locations (e.g., obtained by lateral translation of the aperture) to reduce $\mathcal{S}$ noise. The extent of improvement is limited when compounding images that have some inherent correlation (e.g., caused by insufficient translation of the aperture) [23]. Similarly, each $\hat{R}_{AB}$ contains a new realization of $\mathcal{S}$ noise because each element pair observes the target from a different spatial location. However, this spatial shift is slight, and as with spatial compounding in B-mode imaging, any correlation amongst the correlation estimates would limit the benefits of averaging. The number of *unique* observations of $\mathcal{S}$ is therefore dependent on the correlation of the correlation estimates, a fourth order moment of the backscatter.

III. Proposed Efficient Estimators

We propose three modifications to the currently used spatial coherence estimator in (7): a single sample correlation estimator, a downsampled receive aperture, and an alternative to the averaged estimator.

A. Reduced Axial Kernel

In the first modification, the size of the kernel is reduced to limit negative effects on resolution and to avoid allocating computational resources on efforts that cannot reduce the impact of time-invariant $\mathcal{S}$ noise, as shown above. In the most extreme case, a single sample (0λ) kernel can be used:

$$\hat{R}_{AB}[n] = \frac{a[n]}{|a[n]|} \frac{b[n]}{|b[n]|}. \quad (10)$$

The real component of the single sample estimator can be computed using the in phase and quadrature signals I and Q (omitting n) as

$$\text{Re} \left\{ \hat{R}_{AB} \right\} = \frac{I_a I_b + Q_a Q_b}{\sqrt{(I_a^2 + Q_a^2)(I_b^2 + Q_b^2)}}. \quad (11)$$

Alternatively, for a quasi-monochromatic signal, the complex angle can be used:

$$\text{Re} \left\{ \hat{R}_{AB} \right\} = \cos(\phi_a - \phi_b), \quad (12)$$

where ϕ_a is computed as $\tan^{-1} \frac{Q_a}{I_a}$. An estimate of ρ can be formed subsequently by averaging the correlation estimates across the aperture.

A smaller kernel requires fewer computations, with the single sample formulation reducing the number of computations by a factor of $2T + 1$. The single sample estimate provides the additional benefit that each estimate $\hat{R}_{AB}[n]$ depends only on the signal at n , eliminating cross-dependencies across samples and facilitating parallel processing, such as with GPUs.

B. Downsampled Receive Aperture

In the second modification, a downsampled receive aperture is used to estimate the correlation. If the individual $\hat{R}_{AB}$ correlations that make up $\hat{R}_{\Delta}^{\text{avg}}$ contribute redundant information, an estimate of similar quality should be obtainable using a subset of those correlations. Two simple downsampling techniques are explored: subaperture beamforming (SAB) and uniform downsampling of the aperture.

SAB is a technique in which small segments of the aperture are partially DAS beamformed [32], and can be viewed as applying a spatial anti-aliasing filter followed by a spatial decimation. SAB increases the channel SNR with respect to $\mathcal{T}$ noise by averaging the signals over the subaperture. SAB has been applied to correlation estimation on matrix arrays, demonstrating a substantial reduction in computation time without degrading SLSC imaging performance [17]. SAB has also been used in conjunction with the phase coherence factor to preserve the speckle signal in echocardiography [33].

Uniform downsampling is a sparse sampling of the aperture on receive, where only the signal from every N -th element is retained. This eliminates the need to acquire, focus, or store the majority of the channel signals. Uniform downsampling may be viewed as a spatial decimation without an anti-aliasing filter. As with SAB, uniform downsampling drastically reduces the number of channels, and therefore, computed correlations.

The techniques are illustrated in Fig. 5 along with their corresponding covariance matrices for 4:1 downsampling on a 16 element 1D array. A square in the i -th row and j -th column corresponds to a correlation between the i -th and j -th aperture signal. Only the upper triangle is computed. The diagonal is excluded because the autocorrelation is always equal to one, and the lower triangle is excluded because it is simply the complex conjugate of the upper triangle. In this example, the number of correlations is reduced from 120 to 6. SAB utilizes

the entire aperture, while uniform downsampling uses 1/4th of the array. Downsampling by a factor of D restricts the at which correlations can be estimated to multiples of D , and reduces the number of element pairs in each ξ by D as well.

C. Ensemble Estimator

The third modification changes how the channel pairs are aggregated into $\hat{R}$. The standard technique (as introduced in Sec. II-C) involves computing a separate correlation coefficient for each element pair and reporting the *average* correlation coefficient. Instead, we propose to treat each element pair as a member of a larger ensemble of element pairs with the same population correlation, and to use the whole ensemble to compute a single *ensemble* correlation coefficient. This is more similar to the original definition of the correlation coefficient, defined in (3).

For a single sample kernel, the average correlation is computed as

$$\hat{R}_{\Delta}^{\text{avg}} = \frac{1}{|\xi_{\Delta}|} \sum_{\xi_{\Delta}} \frac{ab^*}{|a||b|}, \quad (13)$$

whereas the ensemble correlation is computed as

$$\hat{R}_{\Delta}^{\text{ens}} = \frac{\sum_{\xi_{\Delta}} ab^*}{\sqrt{(\sum_{\xi_{\Delta}} |a|^2) (\sum_{\xi_{\Delta}} |b|^2)}}. \quad (14)$$

The former estimates the mean of $\hat{R}_{AB}$ for all (A, B) in ξ , which has an expected value of ρ . The latter directly estimates ρ , as per the definition in (3). The average and ensemble correlation estimators have similar computational cost, and can both be used in conjunction with the other proposed efficient techniques as well.

Note that the two estimators are identical when each channel signal has the same magnitude $|a_k| = |b_k| = |s|$:

$$\hat{R}_{\Delta}^{\text{avg}} = \hat{R}_{\Delta}^{\text{ens}} = \frac{1}{|s|^2} \sum_{\xi_{\Delta}} ab^*. \quad (15)$$

However, the magnitude of the signal is a random process with some inherent variance. When the magnitude varies across the aperture, the two estimators apply normalization differently, resulting in different estimates. The differences between the two estimators can be interpreted as follows: the averaged estimator normalizes each correlation estimate separately, whereas the ensemble estimator applies a single normalization to the correlation estimate. The normalization factor for the ensemble estimator is formed using the whole ensemble, and is more stable as a result.

D. Effective Number of Independent Observations

Each estimator can be characterized by estimating the effective number of *independent* observations (K_{eff}) that are obtained. This is similar to a measure proposed for determining the effective number of independent images when compounding B-mode images [34]. That measure assumed a constant magnitude across channel signals to yield a simple formulation. Here, we avoid such assumptions by estimating K_{eff} empirically using simulations. This is accomplished by comparing the variance of the average and ensemble correlation estimators against that of a so-called “best-case” correlation estimator, $\hat{R}_{\Delta}^{\text{best}}$. The best-case estimator is defined as the sample correlation in (3) with the added condition that all K observations are independent and identically distributed (i.i.d.) such that no two observations are correlated and contribute redundant information (i.e., the correlations are themselves uncorrelated).

To compute K_{eff} , the variance of $\hat{R}_{\Delta}^{\text{avg}}$ and $\hat{R}_{\Delta}^{\text{ens}}$ are measured in a uniform homogeneous field of diffuse scatterers. These estimates are composed of $|\xi|$ partially correlated observations. These variances are then compared to that of $\hat{R}_{\Delta}^{\text{best}}$, which is formed using a single element pair with K independent observations. (Independence can be achieved by using a new random realization of scatterers for each observation.) The K that equates the variances is defined as K_{eff} :

$$\sigma^2 [\hat{R}_{\Delta}] = \sigma^2 \left[\frac{\sum_{k=1}^{K_{\text{eff}}} a_k b_k^*}{\sqrt{\sum_{k=1}^{K_{\text{eff}}} |a_k|^2 \sum_{k=1}^{K_{\text{eff}}} |b_k|^2}} \right], \quad (16)$$

where A and B are the single element pair separated by Δ used to compute $\hat{R}_{\Delta}^{\text{best}}$.

This process is illustrated in Fig. 6 with a 16 element transducer to estimate ρ_3 . To estimate $\hat{R}_3^{\text{avg}}$ (top row), all of the received element pairs in ξ_3 are used as observations of ρ_3 for a single scatterer realization. To estimate $\hat{R}_3^{\text{best}}$ (bottom row), only the elements 7 and 10 are used, but over several independent realizations of scatterers. This latter scenario is easily achieved in simulations, where the scatterer positions can be randomly reset.

K_{eff} is a measure of estimator quality. An estimator with high K_{eff} has more stable estimates than one with low K_{eff} . K_{eff} also provides insight into any redundancies in the estimation process. For example, if $\hat{R}_{\Delta}^{\text{avg}}$ is formed using 100 element pairs (i.e. $|\xi| = 100$), but K_{eff} is found to be 5, this would imply that on average, every 20 correlation estimates generates the equivalent of one independent observation of the noise. In this case, it may be possible to obtain the same K_{eff} using a subset of those correlation estimates. An efficient estimator is therefore one that maximizes K_{eff} while minimizing redundant computations.

IV. Methods

A. Simulation Data Acquisition

Field II [35] was used to simulate the Verasonics L12-3v transducer, a 1D uniform linear array with an element pitch of 0.2 mm and an elevation focus of 20 mm. The simulated transducer had 128 elements with a center frequency of 8 MHz and 60% bandwidth. Data was acquired from all 128 channels at a 160 MHz sampling rate. A synthetic transmit aperture was generated using single element transmits to achieve full dynamic focusing on both transmit and receive. Focused channel data were reconstructed in a spatial grid sampled at 1/8th the resolution of the imaging system in each respective dimension.

An ideal speckle target was simulated using a field of homogeneous randomly positioned scatterers. Additionally, 3 mm diameter cylindrical cysts with -12 dB echogenicity at the elevational focus were simulated with the same parameters. To mimic realistic imaging conditions, incoherent acoustical noise was simulated by adding white noise, filtered at the bandwidth of the probe, to the channel signals. Channel SNR was defined to be the ratio the root-mean-square (RMS) of the noise-free channel signals to the RMS of the added noise. In the case of no noise, the channel SNR was ∞ dB, and the only source of estimation error was $\mathcal{S}$ noise.

B. Estimator Performance Metrics

Several quantitative measurements were obtained: estimator variance, K_{eff} , speckle texture SNR, lesion contrast-to-noise-ratio, and computational throughput. Estimator performance was based on the response to the ideal homogeneous speckle target with ∞ dB channel SNR. For each estimator, a spatial coherence estimate $\hat{R}_{\Delta}[n]$ was formed at every reconstructed sample n and for every lag Δ . The variance of the estimator was then computed using all of the samples. However, the variance of correlation estimates grows smaller as ρ approaches ± 1 . To allow for fair comparisons between estimators with different underlying ρ , the variance-stabilizing Fisher transformation [36] was used. The Fisher transformed sample variance (henceforth referred to simply as the variance) is computed as

$$\sigma_{\hat{Z}}^2[\hat{R}_{\Delta}] = \frac{1}{N_{\text{pix}} - 1} \sum_{n=1}^{N_{\text{pix}}} \left(\hat{Z}_{\Delta}[n] - \frac{1}{N_{\text{pix}}} \sum_{m=1}^{N_{\text{pix}}} \hat{Z}_{\Delta}[m] \right)^2, \quad (17)$$

where the $\hat{Z}$ is the Fisher transformed estimate, defined as

$$\hat{Z}_{\Delta} = \frac{1}{2} \ln \frac{1 + \hat{R}_{\Delta}}{1 - \hat{R}_{\Delta}} = \tanh^{-1}(\hat{R}_{\Delta}). \quad (18)$$

The variance was used to compute K_{eff} as follows. First, the variance of the ideal estimator was computed as a function of K :

$$\sigma_{\hat{z},\text{ideal}}^2(\Delta, K) = \sigma_{\hat{z}}^2 \left[\frac{\sum_{k=1}^K a_k b_k^*}{\sqrt{\sum_{k=1}^K |a_k|^2 \sum_{k=1}^K |b_k|^2}} \right], \quad (19)$$

where a_k and b_k were the sets of focused data received from the k -th independent speckle realization on channels A and B with lag Δ , respectively. Only one channel pair (A, B) was used. K_{eff} was then obtained by interpolating $\sigma_{\hat{z},\text{ideal}}^2(\Delta, K)$:

$$K_{\text{eff}} = K : \sigma_{\hat{z},\text{ideal}}^2(\Delta, K) = \sigma_{\hat{z}}^2 \left[\hat{R}_{\Delta} \right]. \quad (20)$$

All computations were performed on an Intel Xeon E5-2687W processor running at 3.1 GHz. Each estimator was implemented as a single-threaded C++ application, and the computation time was taken as the median time over multiple runs. The absolute computational throughput is heavily dependent on a variety of factors such as system configuration, memory bandwidth, and efficient uses of cache. Therefore, these measurements were used only to obtain a rough characterization of how the computations scale (e.g., as a function of window size). Computational throughput was reported as the number of processed pixels per second.

The estimators were also assessed qualitatively with SLSC imaging using the cylindrical cyst simulations. SLSC images were formed by summing the short-lag correlation estimates:

$$V_{\text{SLSC}}[n] = \sum_{\Delta x=1}^{\Delta_{\text{max}}} \hat{R}_{\Delta x}[n], \quad (21)$$

where “short” was defined as $\Delta_{\text{max}} = 16$, corresponding to 1/8th of the aperture. As noted previously, a low variance estimator generates SLSC and CFPD images with a smooth and homogeneous speckle texture, while a high variance estimator will have poor image quality. Image quality was assessed using the texture SNR as measured in a homogeneous region of speckle, and the lesion CNR as measured in a lesion target. The background texture SNR was computed as

$$\text{Texture SNR} = \frac{\mu[\text{Speckle}]}{\sigma[\text{Speckle}]}, \quad (22)$$

where μ_s is the sample mean and σ_s is the standard deviation of V_{SLSC} estimates in a speckle region. A low variance estimator will generate estimates that are smooth and homogeneous, resulting in high texture SNR. A high variance estimator will have erratic estimates, resulting in low texture SNR. Lesion CNR, a measure of lesion detectability [10], [37], was measured as

$$\text{Lesion CNR} = \frac{\mu[\text{Speckle}] - \mu[\text{Lesion}]}{\sqrt{\sigma^2[\text{Speckle}] + \sigma^2[\text{Lesion}]}}. \quad (23)$$

C. Kernel Size Characterization

To understand the impact of kernel size on coherence estimation, correlation estimates were formed using kernel lengths from 0λ up to 1.5λ , in increments of 0.25λ . With the $\lambda/8$ spatial grid in the axial dimension, this corresponded to 1, 3, 5, 7, 9, 11, and 13 samples in the kernel. The texture SNR of the SLSC image was measured for channel SNRs of ∞ dB, +6 dB, and -6 dB as a function of kernel length. The computational throughput was also recorded. Additionally, the full-width half-max (FWHM) of the axial autocorrelation of the speckle texture was measured as a function of kernel length. The FWHM provides a quantification of any visible axial blurring [38]. Finally, the variance of the estimates were computed as a function of lag for several kernel lengths.

D. Downsampled Aperture Characterization

To improve the computational throughput, the receive aperture was downsampled using both the SAB and uniform techniques. For each, the aperture was downsampled by a factor of 2, 4, 8, and 16, corresponding to 64, 32, 16, and 8 channels. For convenience, the type of downsampling will be denoted as “S” for SAB and “U” for uniform, followed by a subscript denoting the downsample factor. For example, U_8 indicates an 8:1 uniform downsampling, while S_4 refers to DAS on 4-element non-overlapping subapertures. The down-sampled estimators were used to construct the $\hat{R}_{\Delta}^{\text{avg}}$ correlation estimates. The mean and variance of the estimates for $\Delta = 16$ were quantified in a uniform speckle target as a function of downsample type and factor, outside and inside the lesion. The kernel size was fixed at 0λ , and the computational throughput was recorded for each estimator.

E. Ensemble Estimator Characterization

The performance of the ensemble estimator ($\hat{R}_{\Delta}^{\text{ens}}$) was compared with that of the averaged estimator ($\hat{R}_{\Delta}^{\text{avg}}$). Correlation estimates were obtained with the two estimators using a kernel size of 0λ . The mean, variance, and K_{eff} of the estimates were measured as a function of lag using either the full aperture or a U_8 downsampled aperture, and with a channel SNR of ∞ dB or +6 dB. Images were formed using DAS, the conventional 1λ $\hat{R}_{\Delta}^{\text{avg}}$ estimator, a 0λ $\hat{R}_{\Delta}^{\text{avg}}$ estimator, a 0λ $\hat{R}_{\Delta}^{\text{ens}}$ estimator, and a 0λ $\hat{R}_{\Delta}^{\text{avg}} U_4$ estimator. The computational throughput was recorded for each method.

F. In Vivo Data Acquisition

A Verasonics Vantage 256 ultrasound scanner was used to acquire *in vivo* echocardiographic channel data in a human subject. Subjects were recruited under the institutional review board protocol Pro00030455 at Duke University. Imaging was performed at the Duke Echocardiography Clinic, and the subject provided written informed consent. Pulse-inversion harmonic imaging was performed using a P4-2v phased array transducer

transmitting at 2 MHz and receiving at 4 MHz. An apical four chamber view of the heart was obtained with a field of view of 72° and 15 cm of depth. The acquired data set was used to make five images: 1) a conventional DAS image, 2) a conventional SLSC image using the averaged estimator and a 1λ kernel, 3) an SLSC image with the averaged estimator and the proposed 0λ kernel, 4) an SLSC image with the proposed ensemble estimator and 0λ kernel, and 5) an SLSC image with the proposed ensemble estimator, 0λ kernel, and U_2 receive aperture downsampling. Here, U_2 downsampling was selected over U_4 because of the greater amount of $\mathcal{T}$ noise present in *in vivo* imaging conditions. The computational throughput of each technique was also measured.

V. Results

A. Reduced Kernel Length

Figure 7a shows SLSC images formed with increasing kernel lengths from left to right. The images were formed with $\tau_{\max} = 16$. The speckle texture in the images grows smoother as the kernel is enlarged. Though the change from image to image is subtle, the image formed using the conventional 1λ kernel shows a significant loss in detail in the speckle as compared to the 0λ case. Lesion CNR values are reported in the top right of the images, and improve with kernel size.

Figure 7b plots the texture SNR as a function of kernel length for three noise conditions: no noise (∞ dB), moderate noise (+6 dB), and heavy noise (−6 dB). The texture SNR improves linearly for all three cases, though the benefit is reduced for the noisy conditions.

The FWHM of the axial auto-correlation of the SLSC image was computed for ∞ dB channel SNR, and is shown in Fig. 7c. The SLSC image is formed with a maximum lag of $\tau_{\max} = 16$. The FWHM of B-mode speckle ($\approx .22$ mm) is also plotted for reference. The FWHM of SLSC increases linearly with the kernel length, and exceeds that of B-mode when the kernel length is $\lambda/2$. The conventional 1λ kernel results in a 92% increase of FWHM over the proposed 0λ kernel.

The computational throughput was measured in pixels computed per millisecond (pixels/ μ s), and is displayed in Fig. 7d. A total of 1, 3, 5, 7, 9, 11, and 13 samples were used for kernels of 0λ , $\lambda/4$, $\lambda/2$, $3\lambda/4$, 1λ , $5\lambda/4$, and $3\lambda/2$, respectively. The proportional increase in throughput relative to the conventional 1λ kernel is listed above each bar. There is a sharp decrease in the throughput as the kernel size increases, with the 1λ kernel being 6.2 times slower than the 0λ kernel.

Figure 7e plots the estimator variance as a function of lag for 4 different kernel sizes. A longer kernel reduces the estimator variance at all lags. This reduction in estimator variance can be linked to the improvements in texture SNR and lesion CNR, both of which are inversely related to the variance. Overall, the kernel improves the estimator variance, texture SNR, and lesion CNR at the cost of axial FWHM and computational throughput.

B. Reduced Sampling of the Aperture

The SAB and uniform downsampling techniques were used to generate SLSC images with the averaged estimator. The images displayed in Fig. 8 were generated using data with a channel SNR of +6 dB and a kernel size of 0λ , and show increasing downsampling from left to right. The SAB downsampled images (Fig. 8a) shows a reduced lesion contrast and a overall brighter texture with increasing downsampling. The uniform downsampled images (Fig. 8b), by contrast, are very similar to the original image, with a slight fine-grained noise becoming apparent at higher downsample factors.

Figure 8c shows the mean correlation estimate values for both downsampled estimators ($N=16$) in the surrounding texture and within the lesion. For the texture estimates, error bars show the variance of the estimates. As seen in the images, the mean value of the SAB estimates increases with higher downsampling, and grows faster inside the lesion than in the texture, leading to reduced contrast. Conversely, the uniformly downsampled estimates are nearly identical to the original, both in mean value and variance.

Figure 8d shows the number of pixels computed per microsecond as a function of downsampling. As expected, downsampling significantly reduces the number of computations required, improving the computational throughput. There appears to be a roughly quadratic increase in computational throughput with respect to downsample factor. The uniform downsampled estimator, which does not have the extra subaperture DAS step, is considerably faster than the SAB estimator at higher downsample factors. With U_4 downsampling, the images are nearly identical to the original, but are formed 13 times faster.

C. Ensemble Estimator

Figure 9a shows reconstructed images of a -12 dB lesion, each labeled with its respective beamformer. The first and second images were both formed using conventional B-mode and SLSC techniques, while the remaining three were formed using progressively more efficient SLSC techniques. The B-mode image displays 40 dB of dynamic range, while the SLSC images show the full positive range of values on a linear scale. The conventional 1λ SLSC image with the averaged estimator has a smooth texture, but on visual inspection appears to have worse resolution than the original B-mode image. In the third image, the proposed 0λ kernel is introduced and improves the averaged estimator SLSC image by eliminating the axial blur. In the fourth image, the averaged estimator is replaced with the ensemble estimator, resulting in a smoother texture without the axial blurring caused by the kernel. In the final image, the U_4 downsampling scheme is applied, generating a slightly noisier version of the fourth image. Fig. 9b shows the computational throughput for the methods used to form each image. As expected, B-mode is several orders of magnitude faster than the more computationally intensive coherence-based techniques. The third, fourth and fifth images in 9a are formed 6, 7, and 105 times faster than the conventional SLSC image, respectively.

Figures 9c–9e compare the averaged estimator to the ensemble estimator. In Figure 9c, the mean values of the texture are shown for noiseless (∞ dB) and noisy (+6 dB) channel conditions as a function of lag. In both cases, the ensemble estimator demonstrate good

correspondence with the averaged estimator. Figure 9d plots the variance of the estimates as a function of lag for the full array (solid lines) and the U_4 downsampled array (markers). For both estimators, downsampling has negligible effect on estimator variance, especially for shorter lags. Furthermore, the ensemble estimator variance is lower than that of the averaged estimator at all lags. Figure 9e plots the number of effective independent observations of $\hat{R}$ that each method contributes, in noiseless and noisy channel conditions. In both cases, the ensemble estimator report higher K_{eff} , suggesting that it utilizes the available information more effectively than the averaged estimator. In the absence of noise, the K_{eff} at $\tau=16$ was 3.3 and 5.3 for the averaged and ensemble estimators, respectively. The overall K_{eff} for both estimators increases with a noisy channel, indicating that the noise contributes to more uncorrelated sample pairs, since the noise is uncorrelated.

D. In Vivo Images

Figure 10 shows *in vivo* images of a human heart. Each image was formed using the same beamforming techniques as in Fig. 9, with the exception of the fifth image, which uses a downsample factor of U_2 instead of U_4 . Fig. 10f shows the computational throughput for each method. The B-mode image in Fig. 10a is computed very quickly (180 times faster than the conventional SLSC image), but much of the apical border of the left ventricle is obscured by high-amplitude clutter. Figures 10b–10e are SLSC images in which the clutter is suppressed. The averaged 1λ image (b) has a visible endocardial border and smooth textures, and appears somewhat blurred throughout. The averaged 0λ image (c) is computed 5 times faster and has a finer grain. The ensemble 0λ image (d) is similar to the averaged image, but with a brighter and more uniform texture appearance in the heart wall. Uniformly downsampling by a factor of two (e) has almost no discernable effect on the image. This final image is computed 20 times faster than the conventional SLSC image.

VI. Discussion

The variance of a spatial coherence estimate can be reduced by repeated observations of a target with independent realizations of noise. We classified noise into two categories in Sec. II: time-invariant ($\mathcal{S}$) and time-dependent ($\mathcal{T}$). By definition, every sample contains a new realization of $\mathcal{T}$ noise, but may or may not contain an independent realization of $\mathcal{S}$ noise. Each observation adds to the overall computational cost. Therefore, it is important to maximize observations of $\mathcal{S}$ noise. We proposed three strategies to improve spatial coherence estimation while minimizing the associated computational cost: a reduced kernel size (deblurring $\mathcal{S}$), a downsampled receive aperture (no impact on $\mathcal{S}$), and an alternative spatial correlation estimator (improving $\mathcal{S}$). The smaller kernel increased computational throughput and reduced axial blurring at the cost of estimator variance and texture SNR; the downsampled receive aperture increased computational throughput with little negative effect, especially when using uniform downsampling; and the ensemble estimator improved the variance of the estimates while marginally improving the computational throughput. These methods were implemented for their strategic sampling of $\mathcal{S}$ noise.

The signal kernel is counterproductive when attempting to reducing $\mathcal{S}$ noise. The kernel increases sampling of $\mathcal{T}$ by increasing the number of available observations. However, the

kernel increases sampling of $\mathcal{S}$ only if the signal is stationary over the axial length of the kernel. Otherwise, the samples within the kernel have different statistics, and their combination manifests as an apparent loss in resolution (Figs. 7a, 7c). Moreover, the extra computations required by the kernel substantially increase the computational cost. In the simulation study, the 1λ kernel (corresponding to 9 samples) was more than 6 times slower than the 1 sample 0λ kernel (Fig. 7d). Considering this significant computational cost, the signal kernel should be applied judiciously, and as an intentional trade of resolution for improved texture SNR. Furthermore, similar effects may be achieved at a lesser computational cost by applying a post-processing filter to the 0λ kernel image. Therefore, from a computational and statistical standpoint, the 0λ kernel (i.e., no kernel) appears to be preferential.

There is a limit to the achievable benefit of averaging correlations with the same expected value. While the averaged estimator is capable of increasing sampling of $\mathcal{S}$ noise, these correlation measurements are highly redundant. Equation (9) shows that the variance of an averaged estimate should improve with more observations if the observations are uncorrelated with one another. This is not seen in Fig. 8c, where the mean and variance of the correlation estimates are unaffected by uniform downsampling, even at a downsampling factor of U_{16} . This lack of change in variance implies that the correlation measurements were highly correlated with one another and that there is no consequence to eliminating these highly redundant measurements. The $\mathcal{S}$ noise could be adequately sampled using a heavily downsampled aperture, and more importantly, at a fraction of the computational cost.

Additionally, the results of this analysis indicate that the ensemble estimator is the most effective way to leverage the wide-sense stationarity of backscatter. The proposed ensemble estimator is similar to the averaged estimator, but uses all of the element pairs to construct a single overall correlation estimate, rather than computing each individual estimate and averaging them together. Though the algorithmic differences are subtle, the ensemble estimator better reflects the original definition of a correlation estimate in (3), and results in a notable improvement in estimator variance (Fig. 9d) in the simulation study. This manifests as a smoother speckle texture in the SLSC images (Fig. 9a). A similar but more subtle effect was observed in the *in vivo* images (Fig. 10). The ensemble estimator is a convenient replacement for the averaged estimator, and is compatible with both the reduced kernel and downsampling methods detailed above, and is also marginally faster to compute (Fig. 9b). The ensemble estimator can be used in place of the averaged estimator to maximize the effectiveness of the available samples.

For a given lag, K_{eff} estimates the effective number of independent realizations of noise that are present in the aperture data. In the noiseless ∞ dB channel SNR environment of the simulation, K_{eff} estimates an upper bound on the number of speckle realizations that a single aperture can observe. In this study, the K_{eff} at short lags was found to be approximately 5 for the ensemble estimator (Fig. 9e). In other words, though there are 112 element pairs at lag 16, the aperture effectively sees 5 independent speckle realizations. This explains why the correlation estimates are highly robust to significant levels of downsampling. In practice, downsampling is not without consequence: in noisy imaging conditions, the 112 element

pairs would observe 112 independent realizations of $\mathcal{T}$ noise (in addition to the 5 realizations of $\mathcal{S}$), and a downsampled aperture could only observe a subset of these. In such cases, SAB downsampling can be used to help improve the channel SNR of each subaperture; however, this method can also lead to a slight increase in variance [17]. Provided that the $\mathcal{T}$ noise levels are low, uniform downsampling is a simple and effective way to quickly obtain correlation estimates that are nearly identical to those provided by the full array.

Combined, these efficient strategies can substantially improve the computational throughput of spatial coherence estimation, making real-time applications of spatial coherence feasible and readily accessible. Though the actual computational speed of coherence estimation will depend on many other factors, such as the hardware and coding implementations, these strategies introduce algorithmic changes that place the computational emphasis where they are most effective in improving spatial coherence estimation based on the physical and statistical limitations of pulse-echo ultrasound imaging.

VII. Conclusion

We have proposed three strategies for efficient spatial coherence estimation: a smaller kernel, a downsampled receive aperture, and an alternative spatial correlation estimator. The kernel was found to significantly increase computation time while causing a blurring of the estimates axially. By eliminating the axial kernel, a 5–6 times improvement in throughput was observed. We found that there was substantial redundancy in the correlation measurements for a given lag. By downsampling the receive aperture uniformly, the computational throughput was dramatically improved with negligible consequence. Finally, the conventional averaged estimator was replaced with an “ensemble” estimator, which better reflects the original formulation of correlation estimation. The ensemble estimator was found to have lower estimator variance than the averaged estimator, even when downsampled. At short lags, the ensemble estimator observed up to 5 effective independent realizations of noise. The three strategies were applied simultaneously to generate simulation images 105 times faster with U_4 downsampling and *in vivo* images 20 times faster with U_2 downsampling.

Acknowledgments

This work is supported by the National Institute of Biomedical Imaging and Bioengineering through grants R01-EB015506 and R01-EB013661. The authors would like to thank the Duke Echocardiography Clinic for their clinical support and Gregg Trahey from Duke University for providing access to imaging equipment for the study.

References

1. O'Donnell M, Flax SW. Phase Aberration Measurements In Medical Ultrasound: Human Studies. *Ultrasonic Imaging*. 1988; 10(1988):1–11. [PubMed: 3291365]
2. Lediju MA, Pihl MJ, Dahl JJ, Trahey GE. Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter. *Ultrasonic Imaging*. Jul; 2008 30(3):151–68. [PubMed: 19149461]
3. Pinton GF, Trahey GE, Dahl JJ. Sources of image degradation in fundamental and harmonic ultrasound imaging: a nonlinear, full-wave, simulation study. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. Jun; 2011 58(6):1272–83.

4. Flax SW, O'Donnell M. Phase-aberration correction using signals from point reflectors and diffuse scatterers: basic principles. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. Jan; 1988 35(6):758–767.
5. Ivancevich NM, Dahl JJ, Smith SW. Comparison of 3-D multi-lag cross- correlation and speckle brightness aberration correction algorithms on static and moving targets. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2009; 56(10):2157–2166.
6. Li PCC, Li ML. Adaptive Imaging Using the Generalized Coherence Factor. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*. Feb; 2003 50(2):128–141.
7. Camacho J, Parrilla M, Fritsch C. Phase Coherence Imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*. May; 2009 56(5):958–74.
8. Synnevag JF, Austeng A, Holm S. Adaptive beamforming applied to medical ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2007; 54(8):1606–1613.
9. Lediju MA, Trahey GE, Byram BC, Dahl JJ. Short-lag spatial coherence of backscattered echoes: imaging characteristics. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. Jul; 2011 58(7):1377–88.
10. Dahl JJ, Hyun D, Lediju M, Trahey GE. Lesion detectability in diagnostic ultrasound with short-lag spatial coherence imaging. *Ultrasonic Imaging*. 2011; 33(2):119–133. [PubMed: 21710827]
11. Jakovljevic M, Trahey GE, Nelson RC, Dahl JJ. In vivo application of short-lag spatial coherence imaging in human liver. *Ultrasound in Medicine & Biology*. Mar; 2013 39(3):534–542. [PubMed: 23347642]
12. Lediju Bell MA, Goswami R, Kisslo JA, Dahl JJ, Trahey GE. Short-Lag Spatial Coherence Imaging of Cardiac Ultrasound Data: Initial Clinical Results. *Ultrasound in Medicine & Biology*. Aug.2013 :1–14. [PubMed: 24210860]
13. Pinton GF, Trahey GE, Dahl JJ. Spatial Coherence in Human Tissue : Implications for Imaging and Measurement. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2014; 61(12):1976–1987.
14. Papadacci C, Tanter M, Pernot M, Fink M. Ultrasound backscatter tensor imaging (BTI): Analysis of the spatial coherence of ultrasonic speckle in anisotropic soft tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2014; 61(6):986–996.
15. Li YL, Dahl JJ. Coherent flow power doppler (CFPD): flow detection using spatial coherence beamforming. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*. 2015; 62(6): 1022–35.
16. Mallart R, Fink M. The van Cittert-Zernike theorem in pulse echo measurements. *The Journal of the Acoustical Society of America*. 1991; 90(5):2718–2727.
17. Hyun D, Trahey G, Jakovljevic M, Dahl J. Short-lag spatial coherence imaging on matrix arrays, Part 1: Beamforming methods and simulation studies. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2014; 61(7):1101–1112.
18. Hyun, D., Trahey, GE., Dahl, JJ. A GPU-based real-time spatial coherence imaging system. In: Bosch, JG., Doyley, MM., editors. *Proceedings of SPIE Vol 8675*. Vol. 8675. Mar. 2013 p. 1B–1-1B–8.
19. Åsen JP, Buskenes JJ, Nilsen CIC, Austeng A, Holm S. Implementing capon beamforming on a GPU for real-time cardiac ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2014; 61(1):76–85.
20. Yiu BYS, Tsang IKH, Yu ACH. GPU-based beamformer: fast realization of plane wave compounding and synthetic aperture imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. Aug; 2011 58(8):1698–705.
21. Lok UW, Li PC. Transform-Based Channel-Data Compression to Improve the Performance of a Real-Time GPU-Based Software Beamformer. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2016; 63(3):369–380.
22. Burckhardt CB. Speckle in ultrasound B-mode scans. *IEEE Transactions on Sonics and Ultrasonics*. Jan; 1978 25(1):1–6.
23. Trahey GE, Allison JW, Smith SW, von Ramm OT. A quantitative approach to speckle reduction via frequency compounding. *Ultrasonic Imaging*. 1986; 164(3):151–164.
24. Carter G. Coherence and time delay estimation. *Proceedings of the IEEE*. 1987; 75(2):236–255.

25. Walker WF, Trahey GE. A Fundamental Limit on Delay Estimation Using Partially Correlated Speckle Signals. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 1995; 42(2):301–308.
26. Kasai C, Namekawa K, Koyano A, Omoto R. Real-Time Two-Dimensional Blood Flow Imaging Using an Autocorrelation Technique. *IEEE Transactions on Sonics and Ultrasonics*. May; 1985 32(3):458–464.
27. Loupas T, Powers J, Gill RW. An Axial Velocity Estimator for Ultrasound Blood Flow Imaging, Based on a Full Evaluation of the Doppler Equation by Means of a Two-Dimensional Autocorrelation Approach. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*. Jul; 1995 42(4):672–688.
28. Walker WF, Trahey GE. Speckle coherence and implications for adaptive imaging. *The Journal of the Acoustical Society of America*. Apr; 1997 101(4):1847–58. [PubMed: 9104014]
29. Goodman, JW. Statistical optics. New York: Wiley-Interscience; 1985.
30. Fedewa RJ, Wallace KD, Holland MR, Jago JR, Ng GC, Rielly MR, Robinson BS, Miller JG. Spatial coherence of the nonlinearly generated second harmonic portion of backscatter for a clinical imaging system. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2003; 50(8):1010–1022.
31. Fedewa RJ, Wallace KD, Holland MR, Jago JR, Ng GC, Rielly MR, Robinson BS, Miller JG. Spatial coherence of backscatter for the nonlinearly produced second harmonic for specific transmit apodizations. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2004; 51(5):576–588.
32. Savord B, Solomon R. Fully Sampled Matrix Transducer for Real Time 3D Ultrasonic Imaging. 2003 IEEE International Ultrasonics Symposium. 2003; 00(c):945–953.
33. Hasegawa H, Kanai H. Effect of subaperture beamforming on phase coherence imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2014; 61(11):1779–1790.
34. Dahl J, Guenther D, Trahey G. Performance Evaluation of Combined Spatial Compounding-Adaptive Imaging: Simulation, Phantom and Clinical Trials. *Ultrasonics, 2003 IEEE 2003; 00(c):1532–1536*.
35. Jensen JA. Field: A Program for Simulating Ultrasound Systems. *Medical & Biological Engineering & Computing*. 1996; 34(1):351–353. [PubMed: 8945858]
36. Fisher RA. On the probable error of a coefficient of correlation deduced from a small sample. *Metron*. 1921; 1(4):1–32.
37. Smith SW, Wagner RF. Ultrasound speckle size and lesion signal to noise ratio: verification of theory. *Ultrasonic imaging*. 1984; 6(2):174–180. [PubMed: 6539979]
38. Wagner RF, Smith SW, Sandrik JM, Lopez H. Statistics of Speckle in Ultrasound B-Scans. *IEEE Transactions on Sonics and Ultrasonics*. May; 1983 30(3):156–163.

Biographies

Dongwoon Hyun was born in Seoul, South Korea, in 1988. He received the B.S.E. degree in biomedical engineering from Duke University, Durham, NC, USA, in 2010, where he is currently pursuing the Ph.D. degree. He is also a Student of New Faculty in bioengineering at Stanford University, Stanford, CA, USA.

His current research interests include beamforming, coherence imaging, and contrast-enhanced ultrasound imaging.

Anna Lisa Crowley received the M.D. degree from the Ohio State University in Columbus, Ohio. She completed Internal Medicine Residency and Cardiology Fellowship training at Duke University, Durham, NC.

She is currently an Associate Professor of Medicine at Duke University and the Director of the Durham VA Echocardiography Laboratory, Durham, NC. Her clinical expertise is cardiovascular imaging of congenital heart disease. Her research interests are optimizing cardiovascular imaging to diagnose and determine the prognosis of patients with congenital heart disease and cardiac infections.

Jeremy J. Dahl (M'11) was born in Ontonagon, Michigan, in 1976. He received the B.S. degree in electrical engineering from the University of Cincinnati, Cincinnati, OH, USA, in 1999, and the Ph.D. degree in biomedical engineering from Duke University, Durham, NC, USA, in 2004.

He is currently an Assistant Professor with the Department of Radiology at Stanford University School of Medicine, Stanford, CA, USA. His research interests include adaptive beamforming, noise in ultrasonic imaging, contrast-enhanced ultrasound imaging, and radiation force imaging methods.

Fig. 1.

The standard correlation estimation process is depicted. K observations of A and B are obtained. The n -th sample of each these are combined to compute a single estimate of $\hat{R}_{AB}[n]$.

Fig. 2.

Correlation estimation with a kernel of 7 samples is depicted. When the number of observations is limited, an axial kernel of samples is often used as a surrogate for observations.

Fig. 3.

The average correlation estimator is depicted for $L = 4$. Each lag 4 channel pair is used to form a $\hat{R}_{AB}[n]$ estimate. These estimates are subsequently averaged together to estimate $\rho_4[n]$.

Fig. 4.

The proposed single sample (0λ) correlation estimator is depicted. $\hat{R}_{AB}[n]$ is estimated using only $a_1[n]$ and $b_1[n]$, dramatically reducing the computational cost.

Fig. 5.

The covariance matrices of a downsampled 16 element 1D array are shown. The sampled correlations are shown as white squares, and the omitted as gray squares. (a) SAB combines elements into large blocks, significantly reducing the number of computed correlations. (b) Uniform downsampling also reduces the number of correlations, but does not sample the full receive aperture.

Fig. 6.

The method of determining K_{eff} is illustrated for a 16 element array. The averaged estimate (top row) is computed for the backscatter from a single acquisition, using all of the lag 3 element pairs in ξ_3 . The best-case estimate (bottom row) is computed using just one element pair, but with multiple acquisitions with independent realizations of random diffuse scatterers. K_{eff} is the number of observations for which the variance of the best-case estimator matches that of the averaged estimator.

Fig. 7.

Spatial coherence estimation with kernels of size 0λ , $\lambda/4$, $\lambda/2$, $3\lambda/4$, 1λ , $5\lambda/4$, and $3\lambda/2$, corresponding to 1, 3, 5, 7, 9, 11, and 13 samples, respectively. (a) SLSC images of a -12dB lesion were formed using increasing kernel sizes from left to right, with a maximum lag of 16. Lesion CNRs are reported in the top right of each image. The (b) texture SNR, (c) axial FWHM, and (d) computational throughput are shown as a function of kernel size. The B-mode FWHM is plotted for reference. (e) The estimator variance as a function of lag is plotted for several kernel sizes. The texture SNR and lesion CNR improves as the kernel size increases, but at the cost of FWHM (92% increase at 1λ) and computational throughput (6.2 times slower at 1λ). The improvements can be attributed to the reduced estimator variance.

Fig. 8.

SLSC images of a -12dB lesion were formed with (a) SAB and (b) uniform downsampling of the receive aperture. (c) The mean estimate values outside and inside the lesion are plotted for $N=16$ as a function of downsample factor. The variance within the texture is shown as error bars. (d) The computational throughput is shown as a function of downsample factor. Overall, both downsampling techniques maintain the same estimator variance while dramatically reducing the computational cost. SAB increases the mean estimated value both inside and outside the lesion, resulting in a loss in contrast. Uniform downsampling yields images that are more similar to the original and is faster than SAB.

Fig. 9.

Ensemble estimator. (a) From left to right, images of a -12dB lesion were formed with DAS (40dB dynamic range), the averaged estimator with a 1λ and with a 0λ kernel, and the ensemble estimator with the full array and with U_4 downsampling. (b) The computational throughput for each of these beamformers is plotted. The speedup over the conventional 1λ kernel is listed above each bar. Plotted as a function of lag are (c) the mean value of the estimates for $\infty \text{ dB}$ and $+6 \text{ dB}$ channel SNR, (d) the estimator variance for the full and uniformly downsampled array, and (e) K_{eff} for $\infty \text{ dB}$ and $+6 \text{ dB}$ channel SNR. The ensemble estimator yields smoother textures than the averaged estimator by reducing the variance of the estimates while maintaining the mean value and computational throughput.

Fig. 10.

In vivo images of the apical four chamber view of the heart are shown for each beamforming technique. (a) A conventional B-mode image, with 40dB of dynamic range, showing clutter near the apex of the heart. (b) SLSC image with the averaged estimator and a 1λ kernel. (c) SLSC image with the averaged estimator and 0λ kernel. (d) SLSC image with the ensemble estimator and a 0λ kernel. (e) SLSC image with the ensemble estimator, 0λ kernel, and U_2 downsampling. (f) The computational throughput for each method is shown. The efficient SLSC image is formed 20 times faster than the conventional.