applied sciences

MDPI

Article

Implementation of Constrained Swept Synthetic Aperture Using a Mechanical Fixture

Nick Bottenus

Department of Mechanical Engineering, University of Colorado Boulder, Boulder, CO 80516, USA; nick.bottenus@colorado.edu

Abstract: Resolution and target detectability in ultrasound imaging are directly tied to the size of the imaging array. This is particularly important for imaging at depth, such as in the detection and diagnosis of hepatocellular carcinoma and other lesions in the liver. Swept synthetic aperture (SSA) imaging has shown promise for building large effective apertures from small physical arrays using motion but has required bulky fixtures and external motion tracking for precise positioning. This study presents an approach that constrains the transducer motion with a simple linear sliding fixture and estimates motion from the ultrasound data itself using either speckle tracking or channel correlation. This work demonstrates, through simulation and phantom experiments, the ability of both techniques to accurately estimate lateral transducer motion and form SSA images with improved resolution and target detectability. In simulation, errors were observed under 83 µm across a 50 mm sweep, and improvements were found of up to 61% in resolution and up to 33% in lesion detectability experimentally even imaging through ex vivo tissue layers. This approach will increase the accessibility of SSA imaging and allow researchers to test its use in clinical settings.

Keywords: beamforming; synthetic aperture; image quality; speckle tracking; displacement estimation

check for updates

Citation: Bottenus, N. Implementation of Constrained Swept Synthetic Aperture Using a Mechanical Fixture. *Appl. Sci.* **2023**, *13*, 4797. https://doi.org/10.3390/app13084797

Academic Editors: Laura Peralta and Kirsten Christensen-Jeffries

Received: 27 February 2023
Revised: 31 March 2023
Accepted: 6 April 2023
Published: 11 April 2023

1. Introduction

Ultrasound imaging is widely used for both screening and diagnosis in soft tissues due to advantages in cost, safety and portability over competing imaging technologies [1–3], particularly with the advent of portable devices [4]. Image contrast is based on differences in the acoustic impedance of tissues, providing good visualization of cysts, tumors, vessels and similar structures. For example, ultrasound has long been the primary screening tool for hepatocellular carcinoma (HCC). However, target detectability in images is based on the combination of contrast and resolution [5]—the latter is a particular weakness of ultrasound imaging. In HCC screening, a recent meta-analysis found that only 45% of early-stage HCCs were detected by ultrasound alone [6], bringing into question its utility without added blood tests or other imaging modalities.

While axial resolution is related to the length of the transmitted imaging pulse (often approximately a wavelength, λ), lateral resolution is based on the wavelength and the f-number, $F = z/D$ for depth z and imaging aperture size D. This leads to an anisotropic resolution that varies drastically through depth. In low-contrast lesions, a loss of resolution blurs lesion boundaries and fills off-axis information from surrounding tissues into the hypoechoic spaces. In abdominal or obstetric imaging, where imaging depths can exceed 10–20 cm, particularly in obese patients, the lateral resolution can be worsened to several millimeters.

This makes the characterization of very small suspected HCC lesions (<1 cm) impossible until further growth has occurred [7], thus, lessening the advantages of early detection. Current LI-RADS protocols suggest the alternatives of contrast-enhanced ultrasound or CT/MRI for the diagnosis and staging of HCC [8,9] with standard B-mode imaging unable to provide clear tumor size measurements.

While the imaging frequency (and wavelength) of ultrasound is limited by the required penetration, with higher attenuation at higher frequencies, the aperture size is fundamentally only limited by the available acoustic window. A standard curvilinear imaging array used in abdominal applications has a total aperture size around 5–6 cm, although these elements point radially outward to increase field of view and result in an even smaller coherent active aperture size. My colleagues and I have recently demonstrated the value of increasing the aperture size to improve the lateral resolution and detectability at depth in simulation [10], ex vivo experiments [11] and in vivo imaging [12] even in the presence of image-degrading clutter effects. Multi-transducer experiments have demonstrated improvements in resolution that scale with the effective aperture extent [13–15] and provided enhanced visualization of in vivo liver and kidney structures.

Despite the promise of large-aperture imaging strategies, there are practical barriers to the implementation of such systems. The cost, weight and electronics complexity of arrays scales with the number of array elements required, particularly when considering 2-D array systems [16]. Point-of-care systems are unlikely to implement large apertures as they are designed to be low cost and highly portable, often imposing data bandwidth constraints on the system as well. Colleagues and I have previously introduced a reduced complexity method for large aperture imaging termed "swept synthetic aperture" (SSA) imaging [12,17] that may be more widely applicable, using the motion of the ultrasound array to form an effective coherent array that is larger than the physical footprint of the device.

This is similar to a monostatic imaging approach [18] with the advantages of improved signal-to-noise ratio and signal redundancy provided by an imaging array. Motion has been extensively used in ultrasound to form larger fields of view [19] and 3-D images [20], both using external tracking systems [21,22] and image-based motion estimation [23,24] approaches; however, these methods simply stitch together B-mode images rather than coherently combining data for higher resolution.

The SSA approach requires precise array position and orientation knowledge during the sweep to perform synthetic aperture beamforming of the data from each array position to a common imaging field of view. It is known that synthetic aperture imaging is particularly sensitive to errors in the axial direction, requiring sub-wavelength precision [25]. SSA imaging similarly requires sub-wavelength jitter error (random position-to-position variation) but is much less sensitive to calibration errors (a fixed error in position/orientation) [12].

Initial SSA investigations found varying impacts of translation and rotation errors in different directions with approximately equal impacts of translation in the axial and lateral dimensions (an achieved reconstruction precision of 0.1 mm given jitter error around $\lambda/5$ or calibration error around 2λ for both dimensions). Previous approaches to SSA imaging have used either a robotic positioning system or a mechanical fixture with external position measurement to achieve this required precision [12,17]. These approaches also required precise spatial calibration to relate the position measurement to the actual position and orientation of the array. These requirements significantly limit the translation of the approach to clinical settings.

This work presents an alternative SSA approach that makes use of the ultrasound array signals themselves to estimate transducer motion with sufficient precision for synthetic aperture beamforming. This approach constrains the motion to one degree of freedom using a simple external fixture with no measurement capabilities to simplify the estimation problem. This work characterizes two methods of estimation—lateral speckle tracking and channel signal correlation—for use in this application. Preliminary phantom results were presented in conference proceedings [26] and are expanded here in addition to a simulation study.

2. Materials and Methods

2.1. Pulse Sequencing

Proper pulse sequencing is essential for successful SSA imaging using motion tracking. The motion tracking methods below assume that there is only slight decorrelation between successive data sets representing a small motion of the array. For a focused transmission to a point in space, a translating array produces decorrelation relative to its length [27]. Increasing decorrelation is associated with increasing variance of displacement estimates [28]. To maximize the correlation, unsteered (0 degree) plane waves were transmitted, and echo data were collected on all array elements. This plane wave can be viewed as a relatively constant transmit field with only small differences near the edges of the array extent.

However, the plane wave transmission is a poor choice for SSA as the field of view only overlaps over a small range of aperture positions [26]. The transmit aperture remains constant (unfocused), and the receive aperture can only effectively double in length due to the finite width of the plane wave. Instead, a diverging wave is used to produce a broad transmit field that shifts with position of the array. This transmit signal creates rapidly decorrelating echoes with translation, a poor choice for displacement estimation but ideal for maximizing the extent of both the effective transmit and receive apertures over a sweep. The effective aperture length reached is calculated as the average of the swept transmit and receive apertures.

These transmit geometries and their trade-offs are illustrated over a sweep of the transducer in Figure 1. Assuming a constant transducer motion that covers 10 cm in 1 s, a pulse repetition frequency of even 1 kHz would provide spatial sampling of 0.1 mm which is oversampled compared to past work [12]. It is expected that, even if the maximum velocity is increased due to acceleration required at the beginning and end of the sweep, the displacement will remain small, on the order of the element pitch. These transmissions can, therefore, be interleaved into a single sequence to provide highly correlated data for displacement estimation and decorrelated data for synthetic aperture imaging as shown in Figure 2.

The simplest sequence would have alternating plane waves and diverging waves that repeat at the chosen repetition frequency. An improved sequence would use two diverging waves, one from either end of the array, with the plane wave. This extends the effective transmit aperture for the SSA image by an entire transducer length (half in each direction at the ends of the sweep). This diagram shows the timing used in this paper with 260 µs between pulses within the sequence and an overall frame rate of 500 Hz. While, in these diagrams, the diverging wave is shown produced by a single element, it is also possible to use subapertures to create diverging virtual sources [29] with an improved signal-to-noise ratio (SNR).

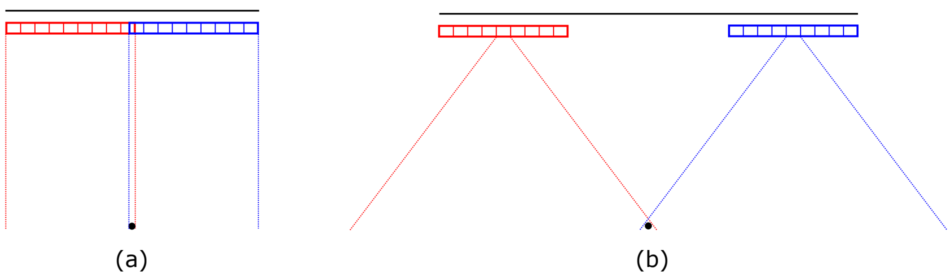


Figure 1. Illustration of the transmit geometries used for correlation-based SSA imaging showing two array positions imaging a target (black circle). Black lines above the array show the synthetic receive array length. (**a**) Plane wave transmission produces highly correlated echoes for motion tracking but limits spatial overlap and does not produce an improved synthetic transmit aperture. (**b**) Diverging wave transmission produces long effective transmit and receive arrays for high resolution due to large spatial overlap at depth but has low echo correlation for motion tracking.

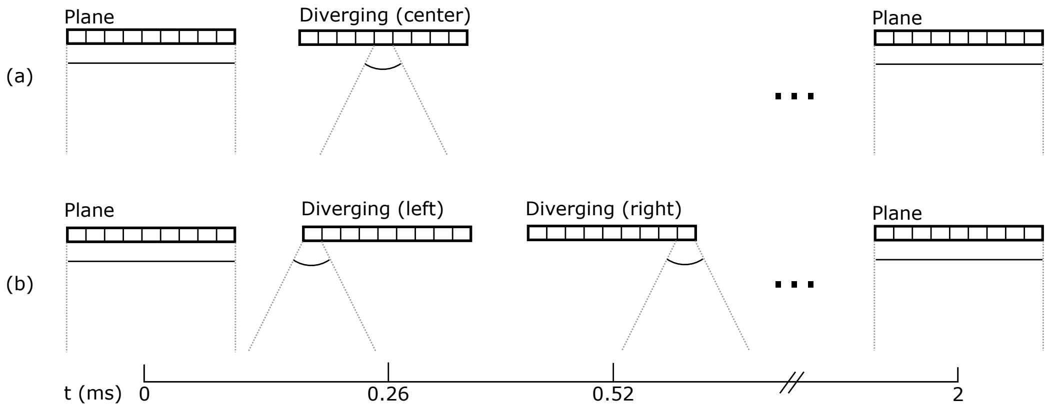


Figure 2. Sample pulse sequences for correlation-based swept synthetic aperture imaging using a 500 Hz frame rate and 260 μs between pulses to allow propagation and data acquisition. The ellipses represent a gap between transmissions. (**a**) Alternating plane and diverging waves with plane waves to provide highly correlated data and diverging waves to provide a wide acceptance angle for coherent combination. (**b**) Diverging waves transmitted from each edge instead of the center to maximize the effective transmit aperture extent.

2.2. Simulation

Field II [30,31] was used to simulate transducer motion for mixed target phantoms to study the optimization of estimation parameters. This simulated the P4-2v phased array transducer with center frequency 3 MHz, 0.3 mm element pitch, 64 elements and 80% bandwidth. The simulations were performed at a 120 MHz sampling rate, and the channel data were stored at a 20 MHz output sampling frequency. A 120 mm (axial) $\times$ 200 mm (lateral) $\times$ 2.5 mm (elevation) speckle-generating material at 15 scatterers per resolution cell was simulated at the expected SSA resolution (approximately equivalent to that of a 7 cm array) with a 4 mm radius anechoic lesion at a 6 cm depth and point target at a 10 cm depth. Simulations were repeated for 10 different realizations of randomly positioned and weighted scatterers. Four sets of simulations were performed:

1. Reference image—Multistatic acquisition (each transmit/receive element pair) from a single array position generates a synthetically focused baseline image without the SSA technique.
2. Subsample estimation sweep—Simulations of plane wave transmission and single-element diverging transmission (from the center of the array) were produced from fine displacements up to a single element pitch in 0.03 mm steps.
3. Large displacement estimation sweep—Simulations of plane wave transmission and single-element diverging transmission were produced from larger displacements as would be found in a 5 cm SSA sweep in steps of 0.1 mm.
4. Equivalent image—Full synthetic aperture acquisition from a fully populated 7 cm long array produces a similar image to the 5 cm SSA approach when the active transmit aperture is limited to 5 cm.

2.3. Phantom Acquisition

2.3.1. Translation Stage

Constrained SSA was achieved with a known ground truth motion profile using a Newport UTM100 linear translation stage (Newport, Irvine, CA, USA). The P4-2v transducer was attached to the stage using a custom molded holder such that the array direction was aligned with the direction of motion. The stage was programmed over a serial communication interface, and motion was synchronized with the imaging pulse sequence described below.

2.3.2. Slide Device

Constrained SSA was achieved without robotic control or external tracking using a 3-D printed slide fixture designed for the P4-2v ultrasound transducer shown in Figure 3. A track was designed allowing up to 10 cm motion (center to center) of the array along a linear path. A sliding block was designed to ride in a cutout track, fitting tightly to prevent tilting or twisting of the array but loose enough to slide freely when moved by hand. A 3-D scan of the P4-2v transducer was used to create a hole the appropriate size and shape in the sliding piece, enlarged to allow for the dimension tolerance of the printed PLA filament. The sliding piece was divided into two pieces to allow for easy insertion and removal of the transducer. The sweep was oriented along the array elements (i.e., within the imaging plane). A handle was added to assist in holding the device still during the transducer sweep but was removable. The pieces were printed using a Lulzbot Taz 5 desktop 3-D printer (Aleph Objects, Loveland, CO, USA).

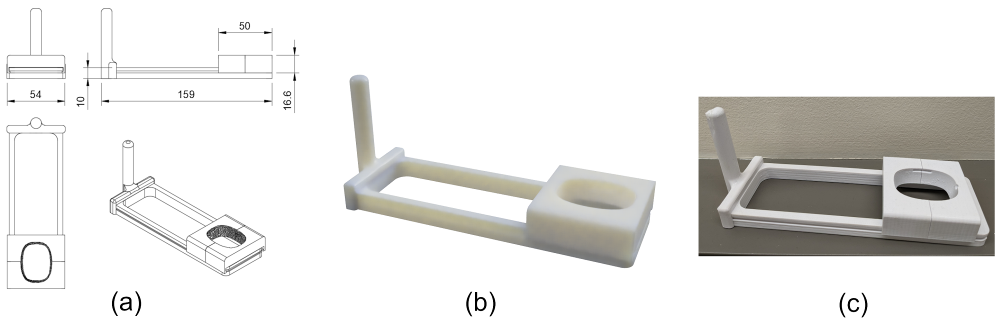


Figure 3. Mechanical slide fixture allowing 1-D movement along the array dimension. (**a**) Dimensioned diagram (units mm). (**b**) Rendering. (**c**) 3-D printed device.

2.3.3. Data Acquisition

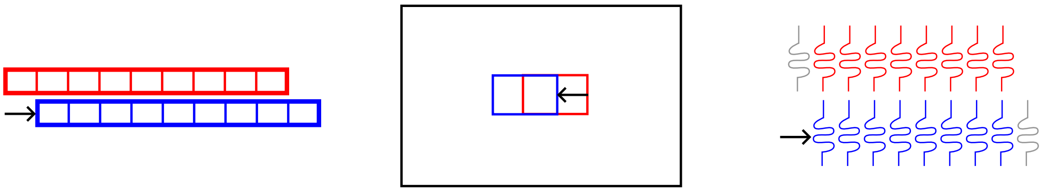
Data were acquired using the Verasonics Vantage 256 research ultrasound scanner (Verasonics Inc., Redmond, WA, USA) and P4-2v phased array transducer. This is a small phased array transducer with 0.3 mm element pitch and 64 elements. Three studies were conducted:

1. The imaging target was the ATS 549 tissue mimicking phantom (CIRS, Norfolk, VA, USA), and the transducer was attached to the translation stage as described above. The transducer was swept over 5 cm in 0.033 mm steps, alternating between plane wave transmission and left/right diverging transmissions (virtual source and 11 elements with f-number −0.75) in that order to mimic a continuous sweep with spacing of 0.1 mm between repetitions.
2. The transducer was used with the linear slide device on the ATS 549 phantom. The phantom was positioned such that point targets were located below lesion targets, and both were visible in the image. The transducer was coupled with gel, and the device was held steady on the phantom surface by hand. The transducer was swept along the track covering approximately 5 cm in 1 s, although the temporal dynamics and extent varied from sweep to sweep. During the sweep, transmissions were alternated as described in Figure 2b at an overall pulse repetition frequency of 500 Hz.
3. The manual sweep was repeated with an ex vivo tissue sample between the transducer and an ATS 539 phantom. The phantom was positioned so that lesion targets were located below point targets, and both were visible in the image. Three different store-bought meats were used to study whether the intervening layer affected motion tracking. Pork chop (12 mm thick), beef loin (10 mm thick) and chicken breast (9–19 mm thick) were used as varying models. Samples were prepared to be wider than the slide device so that they rested entirely on top of the sample. The degassed samples were placed on top of plastic wrap, and all layers were coupled with gel. The same sweep sequence was used as above.

In each case, channel radio frequency (RF) data were stored for offline processing. For each, a reference image (i.e., no sweep) was also acquired using steered plane waves from −45 to 45 degrees.

2.4. Motion Estimation Methods

As transducer motion is constrained to occur within the imaging plane, it should be visible in the imaging data as a contrary motion of the imaging targets. Although ultrasound is best suited to motion estimation in the beam direction [32,33], it is possible to estimate motion in the transverse (lateral) direction using speckle tracking [34] or transverse oscillation [35,36] approaches. An alternative approach is presented below that is well suited to the constrained motion problem. These approaches are shown in Figure 4. In both cases, the same recorded data from plane wave transmissions were used to maintain an approximately constant transmit field and measure the motion of the receiving array (using the diverging wave data only for image formation). While it should be possible to measure displacement from a spatially varying transmit field as well, that would result in additional decorrelation [37] and is outside the scope of this paper.



(a) (b) (c)

Figure 4. (**a**) Assuming low decorrelation over small displacements, motion of the array (here by one element pitch) can be estimated using a cross-correlation search between plane wave transmissions. (**b**) Speckle tracking uses a 1-D search of an image kernel between a reference and target image over a field of view relative to each array position. (**c**) Channel correlation compares shifted pairs of raw channel signals with varying offsets, reducing the active subaperture with increasing shifts.

2.4.1. Speckle Tracking

Speckle tracking uses focused images (radio frequency or envelope detected, channel or beamsum) with cross-correlation to track motion between pairs of frames. It is assumed that the speckle patterns remain largely correlated for small displacements of the receiving array [27] and can be matched. For the lateral estimation, envelope detection was performed on the beamsum data to reduce axial variations and performed all operations on a Cartesian grid.

Given a 2-D reference region (or kernel) from the first plane wave image, normalized cross correlation was used to search laterally in the second plane wave image for the best matching position as shown in Figure 4b. The observed motion in the image is opposite to the motion of the transducer. The parameters of this estimation (kernel size and depth) are explored in the simulation section. For phantom imaging, the kernel and search region were chosen to be 10 × 10 mm at a depth of 20 mm (30 mm for cases with added tissue) with a 2.5 mm lateral search region in each direction.

2.4.2. Channel Correlation

While speckle tracking is effective, it is a tool designed for varying motion estimation throughout a field of view. Here a more direct estimation tool for the problem of transducer displacement is introduced. Channel correlation relies on the idea that, for small motions, there will be subapertures from the initial and displaced position that spatially overlap (possibly with some sub-pitch offset). The backscattered echoes observed by these subapertures should be highly correlated given a relatively constant transmit field. One can, therefore, search for transducer motion by correlating the recorded channel data from these subapertures without the need for focusing or image formation.

Subsets of channels, as shown in Figure 4c, were studied by removing channels from opposing ends of the array for the initial and displaced data sets to increase the search offset. The same estimation depth and axial kernel size were used as in speckle tracking (using the raw echo signals rather than beamformed images) and searched up to offsets of 10 elements (3 mm).

2.4.3. Subsample and Multi-Lag Estimation

For both methods, it is necessary to perform subsample estimation to precisely track motion between frames. For speckle tracking, the beamformed lateral pixel sampling determines the quantization of motion estimates. For channel tracking, estimates were quantized to the element pitch. While it is possible to adjust the choice of lateral pixel sampling to improve the estimation accuracy, that is not possible for the array elements used in channel correlation. Instead, the correlation curve as a function of displacement (lag) was fit to a model allowing us to estimate the location of the peak correlation and the peak correlation value [38]. The iterative reconstructive interpolation method was used for all subsample estimation in this work, although this requires a strictly bandlimited signal and may produce errors for spatially undersampled (aliased) signals. In those cases, a polynomial fit may produce a better result.

In order to reduce both noise and quantization error multi-lag estimation scheme was implemented, combining displacement estimates from multiple combinations of image frames [39]. A measurement matrix $\boldsymbol{H}$ was constructed where each row represents the difference in position between two frames at positions in $\boldsymbol{X}$. The estimated displacements are stored in the matrix $\boldsymbol{\Delta X}$ and may contain some error $\boldsymbol{\epsilon}$:

$$\boldsymbol{HX} = \boldsymbol{\Delta X_{true}} = \boldsymbol{\Delta X} + \boldsymbol{\epsilon} \tag{1}$$

For example, for a sequence with four images at positions X_1–X_4 this becomes:

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} \Delta X_{12} \\ \Delta X_{13} \\ \Delta X_{14} \\ \Delta X_{23} \\ \Delta X_{24} \\ \Delta X_{34} \\ 0 \end{bmatrix} \tag{2}$$

where ΔX_{ij} is the lateral motion estimate produced between the i and j frames. The final row enforces a zero mean condition for the estimates to center the result. A weighted least squares solution to this system was used to estimate the individual frame positions $\hat{\boldsymbol{X}}$, using the peak normalized cross-correlation coefficient in the weighting matrix $\boldsymbol{W}$ to favor highly correlated measurements:

$$\hat{\boldsymbol{X}} = \left(\boldsymbol{H^T W H}\right)^{-\boldsymbol{1}} \boldsymbol{H^T W \Delta X} \tag{3}$$

In this study measurements were limited to a maximum separation of 10 frames, but if computation time were not an issue the $\boldsymbol{\Delta X}$ matrix could be extended with more distant pairs until the correlation coefficient indicated that the estimates were no longer useful.

The code for these methods is made available at https://github.com/bottenuslab/lateral_transducer_tracking (https://doi.org/10.5281/zenodo.7682384).

2.5. Image Reconstruction and Analysis

Reference images (i.e., no sweep) were formed using the plane wave imaging data with standard plane wave focusing methods [40]. Swept synthetic aperture images were formed using standard diverging wave focusing methods with varying transmit and receive aperture positions [12] according to either the translation stage (or simulated) position or the estimated motion (from plane wave transmissions). In both cases, diverging wave data were resampled to retain frames at roughly 0.5 mm spatial intervals, reducing the computation required for image formation and avoiding artifacts due to the varying temporal profiles of the manual sweeps.

It was assumed that all motion occurred along the lateral axis within the imaging plane and that there was no rotation of the transducer. For the interleaved acquisition sequence, the motion estimates from the plane wave frames were interpolated to provide estimates of positions for both the left and right diverging wave transmissions. The focused images from the left and right transmissions were coherently combined, and frames were weighted with a Tukey window to reduce high spatial frequency components of the synthesized transmit aperture present in the SSA image compared to a fully sampled large array [12].

Image analysis was performed using standard methods for resolution and lesion detectability measurement. The lateral point spread function was displayed using a slice through the peak of the point target with the highest value normalized to 0 dB. Lesion detectability was quantified using the generalized contrast-to-noise ratio (gCNR):

$$\mathrm{gCNR} = 1 - \int_{-\infty}^{\infty} \min\{f(x)g(x)\}\, dx, \tag{4}$$

where f and g are the normalized histograms of two image regions [41]. Histogram matching was applied to images across different conditions to visually match the speckle background appearance (the mean and variance) despite variations in the peak value across images [42].

3. Results

3.1. Simulation

3.1.1. Displacement Estimation

Lateral displacement estimation was first investigated using the two proposed techniques—speckle tracking and channel correlation—in simulation. Figure 5 explores the error in estimation faced over varying displacement distances using individual pairs of frames (i.e., no multi-lag estimation). Over short distance scales, up to the element pitch, the two methods demonstrate different behaviors. Speckle tracking (lateral pixel spacing $\lambda/4 = 0.13$ mm) shows an oscillating error possibly related to the acoustic resolution, pixel spacing and the subsample estimation method chosen (iterative reconstruction). It is expected that this behavior could change for varying choices of these parameters. Channel correlation shows a much smoother performance because the channel spacing is much coarser (0.3 mm) and the estimate does not use focused signals. However, it should be noted that, in both cases, the error is limited to less than 5 µm or $\lambda/100$.

Even over larger distances, up to 5 mm in Figure 5, both methods perform well. Channel correlation shows better standard deviation than speckle tracking over this distance, although both are small relative to the wavelength and have a mean near zero. There are also cyclic errors in the estimate here with a period of one element pitch (0.3 mm) for channel correlation. These data suggest that one can use multi-lag estimation representing both small and large displacements to further improve our estimates and can tolerate fairly fast transducer velocities that will lead to several millimeters of displacement between observations.

Challenges that these methods may face in practice, such as noise and sound speed error, were then studied and the best estimation parameters to use determined with the results shown in Figure 6. First, the impact of bandlimited additive Gaussian noise on the estimation was explored. Reducing the channel SNR to 0 dB (measured at the same depth used for displacement estimation) resulted in additional variance in the estimates compared to what was seen in Figure 5, but the estimates remain reasonably unbiased and show standard deviation less than 50 µm over the 5 mm distance. Second, error to the speed of sound used in estimation (including image formation for speckle tracking) was introduced, varying it from 1240 to 1840 m/s compared to the correct 1540 m/s.

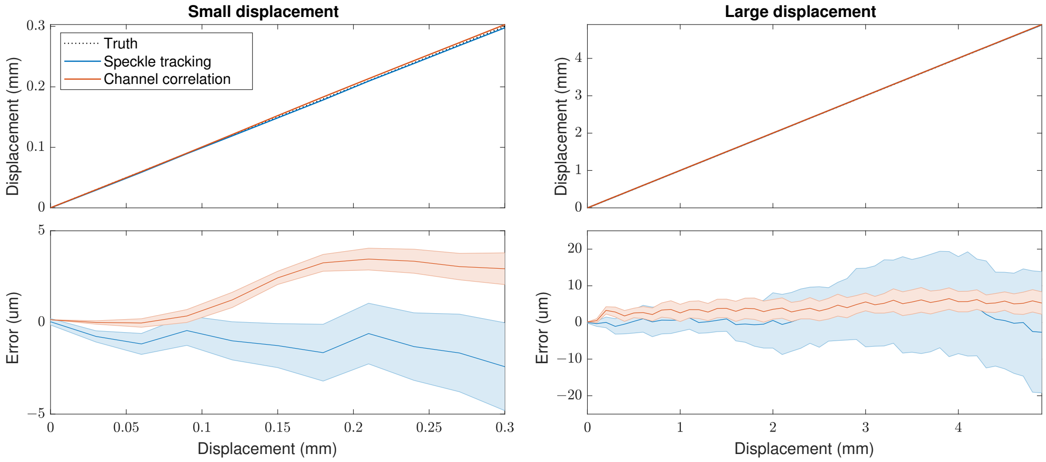


Figure 5. Demonstration of speckle tracking and channel correlation for displacement estimation using individual estimates on pairs of frames. Estimates performed with a 10 × 10 mm kernel. (**Left**) Displacement estimates and error over a short distance, less than one element pitch. Error bars represent 10 speckle realizations. A ±2.5 mm search region was used. (**Right**) Estimates and error over larger distances, up to 5 mm. A ±7.5 mm search region was used. Errors depend on the choice of sampling, kernels and subsample estimation method.

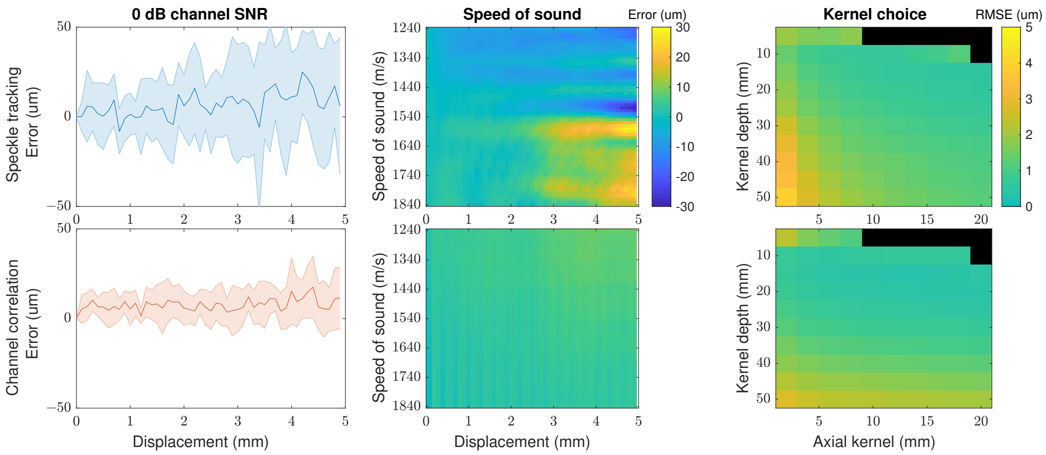


Figure 6. Challenges to and optimization of estimates in simulation. (**Left**) Additive channel noise reduces accuracy compared to Figure 5. (**Middle**) Errors in assumed speed of sound affects speckle tracking accuracy but not channel correlation (true speed of sound 1540 m/s). (**Right**) Speckle tracking is more sensitive than channel correlation to choice of kernel location and depth. (Black indicates regions where the chosen kernel extended beyond the data/image extent.)

This error led to slightly varying depths of estimation and effective kernel sizes but had largely no effect on the channel correlation approach. The speckle tracking approach, due to the distortion of the images produced, showed average positive and negative estimation errors at larger distances up to approximately 30 µm. This bias is likely the result of the stretched speckle pattern used in correlation. Finally, the impact of kernel choice on the estimates was studied. The depth of estimation and the axial kernel extent were varied for both methods and the root mean square error of the estimates were calculated (all showed roughly linear profiles as in previous cases).

Some increase in error at larger depths and reduction of error with a larger axial kernel were observed. Greater sensitivity and larger errors were found for speckle tracking, but both methods had average errors under 5 µm. While better performance was observed for channel correlation across these tests, it should be noted that it is not expected that the errors observed in either case are sufficient to degrade SSA imaging performance.

3.1.2. Imaging Results

The estimates from both methods were used to produce SSA images over a simulated 50 mm transducer sweep of the 19.2 mm array. Figure 7 shows a reference image produced by synthetic aperture imaging (coherent plane wave compounding) from a single transducer position. The full width at half maximum (FWHM) point target resolution at 100 mm depth was measured to be 2.20 mm.

Very similar performance was observed for the two estimation methods compared to image formation using the true applied motion profile, improving the FWHM to 0.70 mm in both speckle tracking and channel correlation, matching the ideal case. The maximum errors in displacement estimation for speckle tracking and channel correlation were 12 and 83 µm, respectively. These results are compared to a fully sampled large array with a 70 mm receive aperture and 50 mm transmit aperture that is roughly equivalent to the effective array formed by SSA and also achieved a FWHM of 0.70 mm.

The observation of some differences due to the spatial frequency content differences between a fully sampled array and SSA effective aperture reflected in a trade-off between resolution and side lobe performance was expected. gCNR was also measured for the anechoic lesion in the images and found similar improvements across the cases with SSA improving gCNR and the estimated motion cases performing close to the ideal cases (slightly lower than with the equivalent large array).

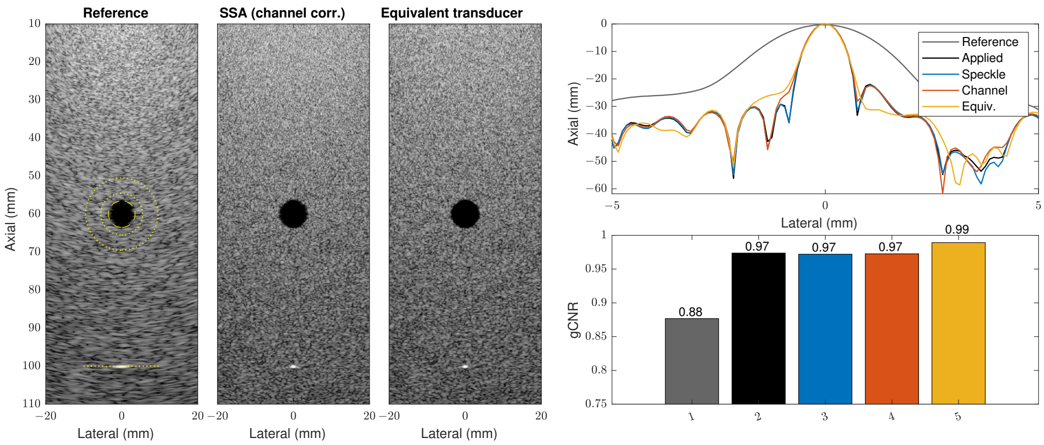


Figure 7. Simulation comparison of reference image, SSA images and equivalent transducer matched to the effective length of the SSA acquisition. (**Left**) Images for reference, channel correlation estimated SSA and equivalent transducer, shown with a 50 dB dynamic range and ROIs indicated for the lesion and point. (**Right top**) Lateral PSF for point at 10 cm depth. (**Right bottom**) gCNR for anechoic lesion target located at 6 cm depth.

3.2. Phantom

3.2.1. Translation Stage

The application of the two estimation techniques to SSA imaging in a series of phantom experiments was investigated. Figure 8 shows the initial experiment comparing the estimates to a known motion as applied by a translation stage. Both methods were able to accurately reproduce the applied motion profile over a span of 50 mm with maximum error for the speckle tracking method around 400 µm and for the channel correlation method under 50 µm. Interestingly, the estimate for channel correlation is biased, not showing the usual linear trend across the sweep. This may be due to the fixed step size creating the same slightly biased error on each step according to the subsample estimation strategy, meaning it would not be seen in freehand sweeps with varying step sizes.

Both strategies produced SSA images comparable to the ideal case. This is further explored in Figure 9 by quantifying the FWHM resolution of a point target and gCNR of a lesion target. The resolution was improved from 3.22 mm in the reference case to 1.04, 0.93 and 1.03 mm in the applied motion, speckle tracking and channel correlation cases, respectively. It should be noted that the PSF for the speckle tracked case noticeably improved from even the applied motion case.

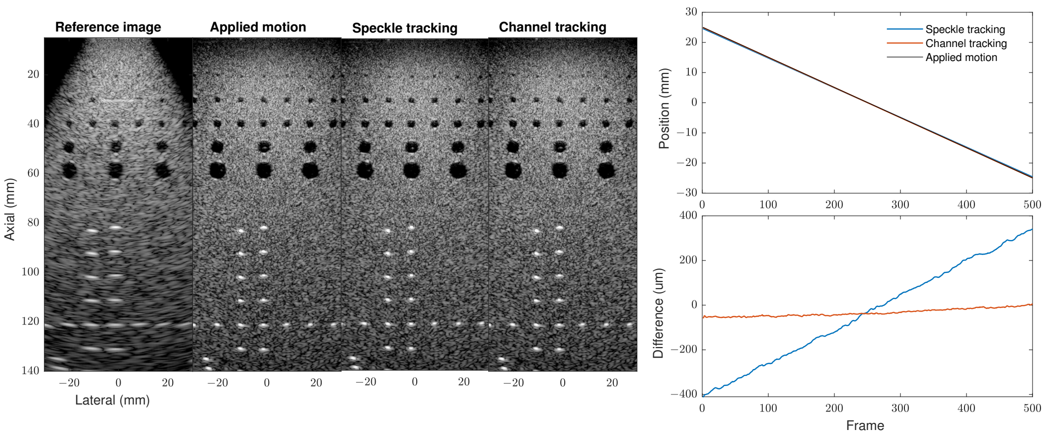


Figure 8. Phantom experiment with lateral motion applied in the translation stage. (**Left**) SSA images show significant resolution improvements compared to the reference image, produced by a 19.2 mm array (50 dB dynamic range). (**Right**) Estimated motion over 50 mm translation shows strong agreement between speckle and channel tracking estimates compared to the applied motion with the error under 400 µm.

It is possible that the translation stage transducer mount was slightly misaligned from the direction of stage motion and that the speckle tracking technique, directly measuring target motion, compensated for this error. This seems plausible as the channel tracking result shows very small improvements over the applied motion case as well, although not as significantly as with the speckle tracking approach. The lesion detectability measured by gCNR similarly confirms these improvements with the speckle tracked SSA case producing the clearest lesion and lesion boundary but all SSA cases improving on the reference case.

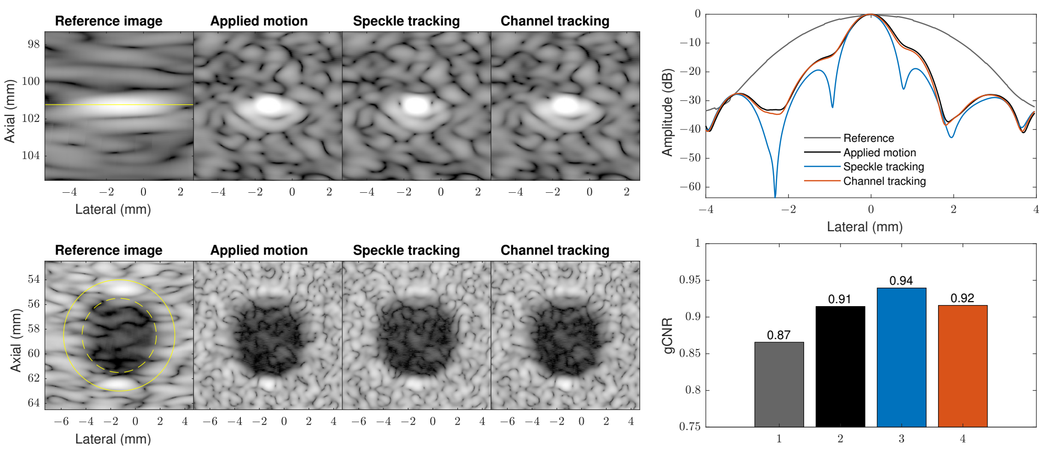


Figure 9. (**Top**) Point target images corresponding to the central point target at depth 100 mm from Figure 8. Yellow line indicates the depth for the lateral PSF. (**Bottom**) Lesion target images corresponding to the central large lesion target at depth 60 mm from Figure 8. Yellow circles indicate ROIs inside the dashed circle and outside the solid circle for gCNR calculation. All images show a 50 dB dynamic range. Both metrics show comparable performance for the two motion tracking techniques and similar performance to the applied motion case with slight improvements using speckle tracking.

3.2.2. Linear Fixture

With an established baseline case showing that channel correlation and speckle tracking produce motion profiles that create SSA images equivalent to the ground truth for an ideal motion case, the application of the channel correlation approach to a freehand sweep was studied using the linear slide fixture. Figure 10 shows four sweeps over approximately the same phantom region with varying temporal profiles and extents due to the manual motion. For image formation, these profiles were resampled using fixed spatial sampling to avoid artifacts from varying sweep speeds, such as would be caused by the stationary region at the end of the yellow curve. Comparable performance was observed across three cases with FWHM of 1.09 (54 mm sweep), 1.07 (49 mm sweep) and 1.14 (48 mm sweep), while the shortest sweep (37 mm) had a slightly larger FWHM of 1.30 mm.

Finally, an additional challenge to the estimation was added by inserting an ex vivo tissue layer under between the transducer and phantom. Figure 11 shows the results of the control case and imaging through the three different tissue layers (pork, beef and chicken). The FWHM improved from 1.47 to 0.59 mm, 1.76 to 0.83 mm, 1.73 to 0.67 mm and 1.94 to 1.39 mm, respectively. The control case had the point target at a smaller depth due to the lack of tissue layer and was expected to, therefore, have lower FWHM and additionally had the longest sweep (72 vs. 41, 51 and 52 mm, respectively).

Qualitatively, the tissue layers were found to add only small amounts of acoustic clutter compared to the control case with the point target in the chicken breast case showing some broadening possibly due to aberration. The gCNR in the reference case was reduced from 0.82 to 0.70, 0.74 and 0.75, respectively, by the tissue layers. These improved to 0.94, 0.93, 0.91 and 0.88, respectively, using SSA imaging.

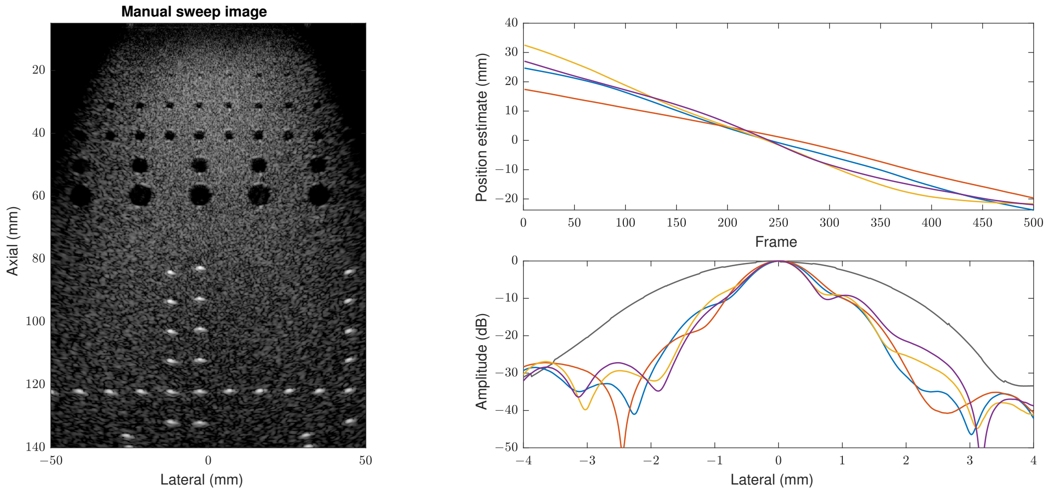


Figure 10. Results of four manual sweeps using linear slide fixture. (**Left**) Sample SSA image from one manual sweep using channel correlation estimate (50 dB dynamic range). (**Right top**) Motion estimates showing variation in sweep extent and temporal profile of manual sweeps. (**Right bottom**) Lateral PSF for manual sweeps showing consistent results with slight variations due to phantom position and sweep extent. Reference PSF from Figure 8 shown in black.

The large difference in FWHM after SSA imaging can be explained by the difficulty of performing the sweep procedure on the chicken breast layer. The tissue was slick, soft and of varying thickness, making it very difficult to hold the fixture still on the surface while sweeping the transducer. The case presented is the best of six acquisitions with the others showing severely degraded images (worse than the reference) due to fixture motion. This difficulty was not observed in the other tissue cases or control case, which all consistently produced high quality SSA results. However, the lesion results showed strong improvement in all cases, even in the chicken breast case despite the limited point target improvement. The gain in both SNR and overall resolution at depth due to SSA imaging greatly improved lesion detectability and image quality.

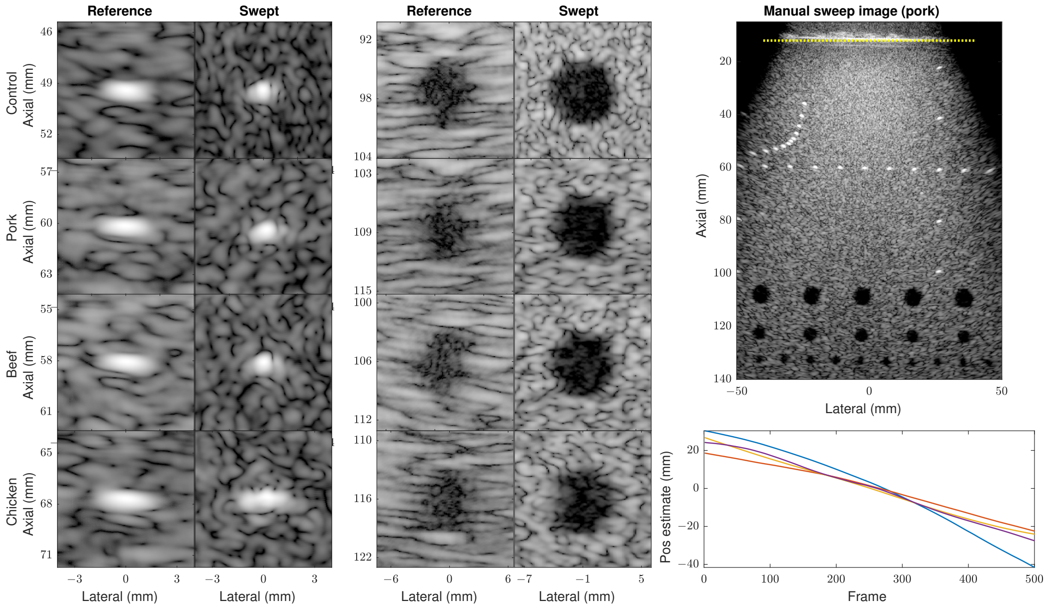


Figure 11. Results of manual sweeps through various tissue phantoms using linear slide fixture. All images show a 50 dB dynamic range. (**Left** and **middle**) Zoomed images of point and lesion targets from reference (no sweep) and manual sweep data. Spatial axes labels are approximate because position varies slightly between sweeps and targets. (**Right top**) Sample SSA image through pork tissue (12 mm thick, above the dotted yellow line). (**Right bottom**) Estimated positions for each case using channel estimation method.

4. Discussion

The simulation and phantom results suggest that both speckle tracking and channel correlation are potential techniques for SSA imaging with a constrained sweep. However, there are distinct trade-offs in these algorithms. Channel tracking showed clear advantages in the accuracy of estimation in these tasks and, importantly, greatly reduced the computational complexity of estimation. For example, in the phantom case of Figure 7, using a Ryzen 7 2700X 8-core CPU, speckle tracking required 8.5 s to beamform the set of 501 image frames (246×164 pixels) used in the correlations and 13 s to perform the lateral cross-correlation search across frame pairs up to 10 frames apart and final weighted least squares estimate.

The channel correlation approach took only 9 s in total (for the cross-correlation search) using the same estimation parameters. In addition to avoiding the need for image formation, channel correlation uses a reduced spatial sampling because it is based on the aperture spacing rather than the sampling requirements in the imaging field. While one could reduce the lateral pixel spacing for speckle tracking closer to the Nyquist limit, the required sampling will depend on the depth of estimation. In both methods, more computation is required for larger kernel and search regions.

The above times should be used for comparison only as GPU beamforming would improve image formation times and a more optimized implementation of normalized cross-correlation (e.g., re-using overlapping calculations and reducing for loops) would improve the correlation times for both methods. The largest challenge in this approach is any motion of the fixture that occurs outside of the imaging plane. These motions cannot be estimated using either a tracking approach or degraded imaging performance as signals from across the sweep cannot be coherently combined. It is demonstrated here that, on relatively firm tissues (pork and beef), the fixture position was able to be maintained with sufficient accuracy to produce SSA images that are comparable to the phantom alone.

Future work will explore the use of these techniques in vivo, where different choices of target may present different scanning challenges. It may be possible, for example, to add an adhesive to the linear fixture if sliding is a problem; however, if the tissue under the fixture compresses, then the fixture may still tilt relative to the anatomy of interest. It may be beneficial to explore coupling the device to skin with a gel pad rather than gel, providing a firmer surface for the fixture and a more consistent coupling across the sweep.

It will also be necessary to study the role that in vivo target movement (e.g., skin deformation, cardiac-induced motion and breathing) has on the SSA images even when the fixture is stationary relative to the skin. A sufficiently rapid sweep relative to the target motion could prevent degradation, or adaptive motion filters may be required to process the data. This study was also limited in its analysis of acoustic clutter, which should be more directly assessed in more challenging imaging cases in the future.

SSA imaging introduces a new dimension of variability in image quality not present in conventional imaging. Even after normalizing for the spatial/temporal dynamics of the sweep, the sweep extent will result in varying effective resolution depending on the length. It may be desirable to limit the sweep length in post-processing to achieve more consistent results between acquisitions.

There are several computational challenges to address for clinical implementation of the constrained SSA approach. First, the current multi-lag displacement estimation approach used in this paper requires all images in the sequence before estimates are made and available for beamforming. A local estimator (or even estimates from individual image pairs) could be used instead so that beamforming could begin during data acquisition for more rapid results. Beamforming itself presents a challenge in that time delays need to be recalculated for each observed array position relative to the target pixel grid.

As a downsampling of acquired frames for beamforming is already implemented, it may be possible to strategically choose these samples to align with precomputed array positions to avoid recalculating delays, although slight position error may result. Finally, over a one second sweep, an operator will be left without live image guidance in the proposed approach. It should be possible to beamform the low-resolution frames for guidance or to interleave a short B-mode imaging sequence with the SSA transmissions to provide live guidance images while the high-resolution SSA image is completed.

If it is possible to make similar estimates in other dimensions, it may be possible to make SSA a truly freehand technique (i.e., without constrained motion). Estimates in the axial dimension are possible as in flow or elasticity imaging, and 2-D motion estimates are becoming more common using vector flow imaging techniques. However, more research is needed to translate these techniques to the problem of estimation transducer motion. Out-of-plane motion will remain a challenge unless using a matrix array, although it is expected that the lower elevation resolution should also reduce the sensitivity to motion in that dimension.

5. Conclusions

This paper demonstrated two signal-processing approaches to transducer motion estimation within a constrained sweep that enabled swept synthetic aperture imaging without external motion tracking. The constrained SSA motion problem is extremely well suited to high-precision estimates given uniform motion of the field within the imaging plane, the ability to use large kernels for estimation and high-frame-rate tracking. This approach may greatly reduce the cost and complexity of an SSA imaging system and is a step towards full freehand motion estimation. The use of this system for in vivo imaging to improve the resolution and target conspicuity at depth without the need for large, expensive imaging arrays will be explored in future work.

Funding: This work was funded by NIH R03-EB032090 from the National Institute of Biomedical Imaging and Bioengineering.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. Code for speckle tracking and channel correlation using the weighted least squares approach, with a sample simulated data set, can be accessed at https://github.com/bottenuslab/lateral_transducer_tracking (https://doi.org/10.5281/zenodo.7682384).

Acknowledgments: The author would like to thank Gregg Trahey, Easha Kuber and Vaibhav Kakkad for contributions to early versions of this work, including 3-D scanning of the P4-2v transducer, speckle tracking code and general discussions. Thanks to Anet Sanchez for assistance in acquiring the ex vivo tissue sweep data.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

FWHM	Full width at half maximum
gCNR	Generalized contrast-to-noise ratio
HCC	Hepatocellular carcinoma
PSF	Point spread function
RF	Radio frequency
ROI	Region of interest
SSA	Swept synthetic aperture

References

1. Royer, D.F. Seeing with Sound: How Ultrasound Is Changing the Way We Look at Anatomy. In *Advances in Experimental Medicine and Biology*; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 47–56. [CrossRef]
2. Miller, D.L.; Abo, A.; Abramowicz, J.S.; Bigelow, T.A.; Dalecki, D.; Dickman, E.; Donlon, J.; Harris, G.; Nomura, J. Diagnostic Ultrasound Safety Review for Point-of-Care Ultrasound Practitioners. *J. Ultrasound Med.* **2020**, *39*, 1069–1084. [CrossRef] [PubMed]
3. Singal, A.G.; El-Serag, H.B. Rational HCC screening approaches for patients with NAFLD. *J. Hepatol.* **2022**, *76*, 195–201. [CrossRef] [PubMed]
4. European Society of Radiology (ESR). ESR statement on portable ultrasound devices. *Insights Imaging* **2019**, *10*, 89. [CrossRef]
5. Smith, S.W.; Wagner, R.F.; Sandrik, J.F.M.F.; Lopez, H. Low contrast detectability and contrast/detail analysis in medical ultrasound. *IEEE Trans. Sonics Ultrason.* **1983**, *3*, 164–173. [CrossRef]
6. Tzartzeva, K.; Obi, J.; Rich, N.E.; Parikh, N.D.; Marrero, J.A.; Yopp, A.; Waljee, A.K.; Singal, A.G. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. *Gastroenterology* **2018**, *154*, 1706–1718.e1. [CrossRef]
7. Dong, Y.; Teufel, A.; Wang, W.P.; Dietrich, C.F. Current Opinion about Hepatocellular Carcinoma <10 mm. *Digestion* **2021**, *102*, 335–341. [CrossRef]
8. Eisenbrey, J.R.; Gabriel, H.; Savsani, E.; Lyshchik, A. Contrast-enhanced ultrasound (CEUS) in HCC diagnosis and assessment of tumor response to locoregional therapies. *Abdom. Radiol.* **2021**, *46*, 3579–3595. [CrossRef]
9. American College of Radiology. US LI-RADS v2017 CORE. Available online: https://www.acr.org/-/media/ACR/Files/RADS/LI-RADS/LI-RADS-US-Algorithm-Portrait-2017.pdf (accessed on 10 April 2023).
10. Bottenus, N.; Pinton, G.; Trahey, G. The Impact of Acoustic Clutter on Large Array Abdominal Imaging. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2020**, *67*, 703–714. [CrossRef]
11. Bottenus, N.; Long, W.; Morgan, M.; Trahey, G. Evaluation of Large-Aperture Imaging Through the ex Vivo Human Abdominal Wall. *Ultrasound Med. Biol.* **2018**, *44*, 687–701. [CrossRef]
12. Bottenus, N.; Long, W.; Zhang, H.K.; Jakovljevic, M.; Bradway, D.P.; Boctor, E.M.; Trahey, G.E. Feasibility of Swept Synthetic Aperture Ultrasound Imaging. *IEEE Trans. Med. Imaging* **2016**, *35*, 1676–1685. [CrossRef]
13. Peralta, L.; Gomez, A.; Luan, Y.; Kim, B.; Hajnal, J.V.; Eckersley, R.J. Coherent Multi-Transducer Ultrasound Imaging. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2019**, *66*, 1316–1330. [CrossRef]
14. Peralta, L.; Ramalli, A.; Reinwald, M.; Eckersley, R.J.; Hajnal, J.V. Impact of Aperture, Depth, and Acoustic Clutter on the Performance of Coherent Multi-Transducer Ultrasound Imaging. *Appl. Sci.* **2020**, *10*, 7655. [CrossRef]
15. Foiret, J.; Cai, X.; Bendjador, H.; Park, E.Y.; Kamaya, A.; Ferrara, K.W. Improving plane wave ultrasound imaging through real-time beamformation across multiple arrays. *Sci. Rep.* **2022**, *12*, 13386. [CrossRef]

16. Wodnicki, R.; Kang, H.; Li, D.; Stephens, D.N.; Jung, H.; Sun, Y.; Chen, R.; Jiang, L.M.; Cabrera-Munoz, N.E.; Foiret, J.; et al. Highly Integrated Multiplexing and Buffering Electronics for Large Aperture Ultrasonic Arrays. *BME Front.* **2022**, *2022*, 9870386. [CrossRef]
17. Zhang, H.K.; Cheng, A.; Bottenus, N.; Guo, X.; Trahey, G.E.; Boctor, E.M. Synthetic tracked aperture ultrasound imaging: Design, simulation, and experimental evaluation. *J. Med. Imaging* **2016**, *3*, 027001. [CrossRef]
18. Ylitalo, J.; Ermert, H. Ultrasound synthetic aperture imaging: Monostatic approach. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1994**, *41*, 333–339. [CrossRef]
19. Weng, L.; Tirumalai, A.; Lowery, C. US extended-field-of-view imaging technology. *Radiology* **1997**, *203*, 877–880. [CrossRef]
20. Prager, R.W.; Ijaz, U.Z.; Gee, A.H.; Treece, G.M.; Wells, P.N.T. Three-dimensional ultrasound imaging. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **2010**, *224*, 193–223. [CrossRef]
21. Prager, R.W.; Gee, A.; Berman, L. Stradx: Real-time acquisition and visualization of freehand three-dimensional ultrasound. *Med. Image Anal.* **1999**, *3*, 129–140. [CrossRef]
22. Chung, S.W.; Shih, C.C.; Huang, C.C. Freehand three-dimensional ultrasound imaging of carotid artery using motion tracking technology. *Ultrasonics* **2017**, *74*, 11–20. [CrossRef]
23. Zheng, S.; Huang, Q.; Jin, L.; Wei, G. Real-time extended-field-of-view ultrasound based on a standard PC. *Appl. Acoust.* **2012**, *73*, 423–432. [CrossRef]
24. Prevost, R.; Salehi, M.; Jagoda, S.; Kumar, N.; Sprung, J.; Ladikos, A.; Bauer, R.; Zettinig, O.; Wein, W. 3D freehand ultrasound without external tracking using deep learning. *Med. Image Anal.* **2018**, *48*, 187–202. [CrossRef] [PubMed]
25. Karaman, M.; Bilge, H.; O'Donnell, M. Adaptive multi-element synthetic aperture imaging with motion and phase aberration correction. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1998**, *45*, 1077–1087. [CrossRef] [PubMed]
26. Bottenus, N. Estimation of transducer translation using channel-domain correlation. In Proceedings of the 2019 IEEE International Ultrasonics Symposium (IUS), Glasgow, Scotland, 6–9 October 2019; pp. 1009–1012. [CrossRef]
27. O'Donnell, M.; Silverstein, S. Optimum displacement for compound image generation in medical ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1987**, *35*, 470–476. [CrossRef] [PubMed]
28. Walker, W.F.; Trahey, G.E. Fundamental limit on delay estimation using partially correlated speckle signals. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1995**, *42*, 301–308. [CrossRef]
29. Karaman, M.; O'Donnell, M. Synthetic aperture imaging for small scale systems. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1995**, *42*, 429–442. [CrossRef]
30. Jensen, J.A.; Svendsen, N.B. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1992**, *39*, 262–267. [CrossRef]
31. Jensen, J.A. Field: A Program for Simulating Ultrasound Systems. *Med. Biol. Eng. Comput.* **1996**, *34*, 351– 353.
32. Kasai, C.; Namekawa, K.; Koyano, A.; Omoto, R. Real-Time Two-Dimensional Blood Flow Imaging Using an Autocorrelation Technique. *IEEE Trans. Sonics Ultrason.* **1985**, *32*, 458–464. [CrossRef]
33. Loupas, T.; Powers, J.; Gill, R.W. 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 Trans. Ultrason. Ferroelectr. Freq. Control* **1995**, *42*, 672–688. [CrossRef]
34. Trahey, G.E.; Allison, J.W.; von Ramm, O.T. Angle independent ultrasonic detection of blood flow. *IEEE Trans. Biomed. Eng.* **1987**, *34*, 965–967. [CrossRef]
35. Jensen, J.A.; Munk, P. A new method for estimation of velocity vectors. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1998**, *45*, 837–851. [CrossRef]
36. Anderson, M.E. Multi-dimensional velocity estimation with ultrasound using spatial quadrature. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1998**, *45*, 852–861. [CrossRef]
37. Trahey, G.; Smith, S.; von Ramm, O. Speckle Pattern Correlation with Lateral Aperture Translation: Experimental Results and Implications for Spatial Compounding. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **1986**, *33*, 257–264. [CrossRef]
38. Céspedes, I.; Huang, Y.; Ophir, J.; Spratt, S. Methods for Estimation of Subsample Time Delays of Digitized Echo Signals. *Ultrason. Imaging* **1995**, *17*, 142–171. [CrossRef]
39. VanDecar, J.; Crosson, R. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares. *Bull. Seismol. Soc. Am.* **1990**, *80*, 150–169.
40. Montaldo, G.; Tanter, M.; Bercoff, J.; Benech, N.; Fink, M. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2009**, *56*, 489–506. [CrossRef]
41. Rodriguez-Molares, A.; Marius, O.; Rindal, H.; Jan, D. The Generalized Contrast-to-Noise ratio: A formal definition for lesion detectability. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2019**, *67*, 745–759. [CrossRef]
42. Bottenus, N.; Byram, B.; Hyun, D. Histogram matching for visual ultrasound image comparison. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2020**, *68*, 1487–1495. [CrossRef]