Accelerated 2D Real-Time Refraction-Corrected Transcranial Ultrasound Imaging

Mozaffarzadeh, M.; Verschuur, D.J.; Verweij, M.D.; de Jong, N.; Renaud, G.G.J.

DOI 10.1109/TUFFC.2022.3189600
Publication date 2022
Document Version Accepted author manuscript
Published in IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control


Important note To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
Accelerated 2D Real-Time Refraction-Corrected Transcranial Ultrasound Imaging

Moein Mozaffarzadeh*, Eric Verschuur, Martin D. Verweij, Nico de Jong, Guillaume Renaud

Abstract—In a recent study, we proposed a technique to correct aberration caused by the skull and reconstruct a transcranial B-mode image with a refraction-corrected synthetic aperture imaging scheme. Given a sound speed map, the arrival times were calculated using a fast marching technique (FMT), which solves the Eikonal equation and therefore is computationally expensive for real-time imaging. In this paper, we introduce a two-point ray tracing method, based on Fermat’s principle, for fast calculation of the travel times in the presence of a layered aberrator in front of the ultrasound probe. The ray tracing method along with the reconstruction technique are implemented on a graphical processing unite (GPU). The point spread function (PSF) in a wire phantom image reconstructed with the FMT and the GPU-implementation was studied with numerical synthetic data and experiments with a bone-mimicking plate and a sagittally-cut human skull. The numerical analysis showed that the error on travel-times is less than 10% of the ultrasound temporal period at 2.5MHz. As a result, the lateral resolution was not significantly degraded compared with images reconstructed with FMT-calculated travel times. The results using the synthetic, bone-mimicking plate and skull dataset showed that the GPU-implementation causes a lateral/axial localization error of 0.10/0.20 mm, 0.15/0.13 mm and 0.26/0.32 mm compared to a reference measurement (no aberrator in front of the ultrasound probe), respectively. For an imaging depth of 70 mm, the proposed GPU-implementation allows reconstructing 19 frames per second with full synthetic aperture (96 transmission events) and 32 frames per second with multi-angle plane wave imaging schemes (with 11 steering angles) for a pixel size of 200 µm. Finally refraction-corrected power Doppler imaging is demonstrated with a string phantom and a bone-mimicking plate placed between the probe and the moving string. The proposed approach achieves a suitable frame rate for clinical scanning while maintaining the image quality.

Index Terms—Transcranial ultrasound imaging; Graphical processing unite, Adaptive beamforming; Phase aberration correction; Temporal bone.

I. INTRODUCTION

Real time ultrasound imaging of the brain started in the late sixties with the development of the first electronic two-dimensional scanner developed by Somer [1]. Decades after, transcranial ultrasound imaging (TUI) remains very challenging and offers poor image quality compared to ultrasound imaging of many body regions where no ultrasound propagation through bone occurs. Poor image quality of TUI results from strong wave aberration [2] and multiple scattering caused by the skull [3, 4]. Current commercial TUI devices ignore the skull.

TUI is nowadays available in most hospitals, clinics and emergency medicine services (EMS) worldwide [5-8]. Compared to CT and MRI, ultrasound imaging is safe (non-ionizing), portable (can be used at bed-site and EMS) and relatively inexpensive. Despite its poor image quality, TUI was shown to be useful for the diagnosis of stroke [8-14], prevention of stroke in children (between an age of 2 and 16) with sickle cell disease [15-17], and vasospasm detection after subarachnoid hemorrhage [18, 19].

TUI is often performed through the temporal window (the thinnest part of the skull that gives the most optimal ultrasound access to the brain) where the squamous part of the temporal bone often consists of a single layer of cortical bone [20-23]. Phase aberration correction was first proposed by modeling the temporal bone as an infinitesimally thin aberrating layer at the surface of the transducer (so called the near-field phase-screen aberration model) [24-26], but the correction obtained by this approach is limited (axially and laterally) to a certain region called the isoplanatic patch [3, 27-30]. Another approach is using either ultrasound measurements [31-33] or CT/MRI scans of the skull to obtain the true geometry and sound speed of the temporal bone prior to image reconstruction, and then correct phase aberration and refraction during image reconstruction [34-37]. Recently, we have shown the feasibility of single-sided two-dimensional transcranial ultrasound through the human temporal window using a single handheld commercial probe, where the position, true geometry and sound speed of the bone layer were estimated for an accurate correction of phase aberration and refraction [38, 39]. The same methodology was used before for in vivo imaging of the inner structure of the...
radius and tibia bone [40, 41]. While promising results were achieved, no fast-enough implementation is available for real-time transcranial imaging, i.e. enabling an imaging rate of at least 20 images per second, which limits its application in practice, especially for translation to 3D TUI [38, 42-44]. Real-time image reconstruction is essential in practice as it allows the operator to freely move the probe find the proper imaging window and imaging plane, and look for biomarkers related to a possible brain disorder in either B-mode or flow images. A frame rate of 4 Hz was reported in [41], but still not fast enough for real-time visualization.

Over the past two decades, the emergence of graphics processing units (GPUs) has facilitated high performance computing [45-47]. Medical ultrasound imaging also took advantage of this technology [48-50]. Different beamforming techniques such as Capon [49, 51-53], short-lag spatial coherence [50, 54-56], synthetic aperture sequential beamforming [57, 58], delay-multiply-and-sum [59] and double-stage delay-multiply-and-sum [60, 61], and volumetric imaging systems [62-64] were accelerated by GPU. GPU-based beamforming softwares were developed [65, 66], and ultrasound vector flow imaging systems were accelerated [67-69]. Yiu, et al. [70] reported on a very high frame rate realization of the synthetic aperture and multi-angle plane wave imaging schemes with 2 GPUs and a recursive implementation, which can output between 1000 and 5000 frames per second depending on the number of receive channels and image depth.

Despite all the acceleration achieved by GPUs in medical ultrasound imaging field, implementations are all based on a strong assumption: an imaging medium with a constant wave-speed (~1540 m/s). This does not apply to TUI. In this paper, we introduce an accelerated image reconstruction technique that corrects for phase aberration and wave refraction caused by the human temporal bone for real time imaging. The two-point ray tracing concept (introduced by Waltham [44]) was used to find the shortest (following Fermat’s principle) travel-time connecting an image pixel to an array element or a virtual point source (for diverging or plane wave imaging) [71]. Unlike in [40, 41], the approach is not iterative and the near and far surfaces of the bone layer are described in a discrete manner using the image grid points. This allows a faster calculation of the travel-times and therefore tremendously reduces computational time for image reconstruction. Its implementation on a GPU further reduces computational time.

II. ACCELERATED RAY-TRACING-BASED RECONSTRUCTION

A. Principle of the proposed method

Image reconstruction is most often performed with a delay-and-sum algorithm, which requires the accurate calculation of the ultrasound round-trip travel times (i.e. from the source to the image pixel and from the image pixel back to the array elements). Several techniques exist to calculate these travel times in a layered medium. Eikonal solvers propagate the wavefront through the medium to calculate the travel-time from a array element or a virtual point source to an image pixel in the brain region [38, 41-43]. Another approach is iterative two-point ray tracing, where the ray path providing the minimum travel time (Fermat’s principle) is searched iteratively and accurately using a method for function minimization, for instance a ray bending approach using Brent’s algorithm [39-41]. However, eikonal solvers and iterative two-point ray

---

Figure 1. A summary of the proposed refraction-corrected image reconstruction approach. Three steps of ray tracing and adaptive beamforming are used for reconstructing the image in the brain area; the illustration is shown only for one pixel (white square) and one element (the element colored green). The rays pass through intermediate points (IPs). A limited number of IPs and travel-paths are shown for simplicity. Red, black and purple squares are IPs defined on the surface of the silicone rubber front layer in the transducer (SIPs), near (NIPs) and far (FIPs) surfaces of the bone layer, respectively. The model shown in (c) presents the numerical model used for imaging in our previous [38] and this study.
tracing are too computationally expensive to reconstruct more than 20 images per second. In contrast, our approach is less accurate but much faster because it seeks the shortest travel-times by evaluating the travel time for a limited number of possible ray paths passing through a limited number of intermediate points (IPs). IPs are defined as the image pixels describing the near and far surfaces of the bone layer in the image. The terms "near surface" and "far surface" are defined with respect to the probe. A summary of the proposed accelerated ray-tracing based imaging approach is provided in Figure 1. We propose the following three steps to reconstruct a refraction-corrected image:

1. The pixels on the far surface of the silicone rubber front layer in the transducer are defined as intermediate points (SIPs; the red squares in Figure 1). The ray tracing starts to find the shortest travel-time from pixels in the skin (up to a certain maximum depth where we expect to detect the near surface of the bone, see the red dashed line in Figure 1(a)) to each array element through the SIPs; see the red vectors in Figure 1(a). The sound speed of the skin should be used to convert the distance to travel-time. The image is reconstructed within the skin area up to the maximum expected depth of the near surface of the skull and using the estimated travel-times. Thus, the image contains the near surface of the bone layer, which is then segmented using Dijkstra’s algorithm [72]. This algorithm seeks the path that crosses the image from left to right and follows the image pixels with the highest intensity in the ultrasound image, by maximizing a merit; here the sum of the pixel values along the path is used.

2. The IPs are updated (called NIPs) to the pixels on the segmented near surface (see the black squares in Figure 1(b)). The ray tracing finds the shortest travel-time from pixels in the bone (up to a certain maximum depth where we expect to detect the far surface of the bone (see the black dashed line in Figure 1(b)) to each array element through the NIPs; the black vectors in Figure 1(b) are examples of the paths. The compressional wave speed of the bone should be used to convert the distances to travel-times. The image is reconstructed up to the maximum expected depth of the far surface of the skull and using the estimated travel-times. The far surface of the bone layer is segmented using Dijkstra’s algorithm.

3. The IPs are updated along the far surface of the bone layer (called FIPs; see the purple squares in Figure 1(c)). Finally, the ray tracing finds the shortest travel-time from pixels in the brain (up to the maximum imaging depth) to each array element through the FIPs (the purple squares in Figure 1(c)), and the rest of the image is reconstructed.

We use prior knowledge of the silicone rubber front layer thickness and wave speed in the silicone rubber front layer in step 1. We also assume that the wave speed in the layer between the probe and the aberrator and the wave speed in bone is known. A value from the literature [73] can be used or it can be estimated with the bidirectional head-wave technique and an autofocus approach as described in [40, 74, 75]. Describing locally the skull as a single homogeneous layer is only valid for a temporal bone without a diploe. Moreover we hypothesize negligible wave-speed anisotropy in bone, in the image plane. The depths of the first and second reconstruction steps (the red and black dashed lines, respectively, see Figure 1) are determined by prior knowledge on the thickness of the skin and the temporal bone [20, 76]; the temporal bone thickness of the adult human skull was measured 2.5 ± 0.9 mm with CT [20]. Results obtained by the proposed reconstruction method are titled/captioned ARC (standing for Accelerated Refraction Correction technique) throughout the paper. While the near surface of the bone layer is depicted flat in Figure 1, our approach can be applied to any irregular interface (such as the far surface of the bone layer in Figure 1).
\[ N_{\text{grid}} = \left\{ \begin{array}{ll} \frac{N_{\text{OP}}}{\text{block} \times X}, & \text{if } N_{\text{OP}} > 0 \vphantom{\frac{N_{\text{OP}}}{\text{block} \times Y}} \\ \frac{N_{\text{OP}}}{\text{block} \times Y} + 1, & \text{else} \end{array} \right. \]
A 4.2 mm-thick bone-mimicking plate (Sawbones, Pacific Research Laboratory, Inc., Vashon, WA) (see Figure 2(b) of [38]) having attenuation close to that in low-porosity cortical bone, a compressional wave speed in the image plane of 3000 m/s and mass density of 1640 kg/m³ [41] was used in the first experiment. A sagittally-cut human skull (see Figure 2(c) of [38]) was used in the second experiment. The temporal bone of the skull was used as the imaging window. The wave speed in the bone layer and water was considered 3500 m/s and 1500 m/s, respectively. We imaged multiple wires having a diameter of 50 µm with a one-dimensional phased array probe (P4-1, ATL/Philips, 2.5 MHz, 96 elements, pitch = 0.295 mm). To reconstruct power Doppler images, a moving string was imaged with a multi-angle plane wave imaging sequence (5 plane waves in -15:15 degrees in water) with (see Figure 3(b)) and without (to acquire a reference dataset) the Sawbones plate in front of the probe; the near surface of the Sawbones plate was at the depth of 5.7 mm. An ensemble length of 100 frames was used to compute the power Doppler images. Singular value decomposition (SVD) filtering was used to remove the stationary signals; Matlab “svd” command was utilized and the first 10 singular values were removed. Spatio-temporal SVD filtering generally performs better than a simple temporal high pass filter for power Doppler imaging, especially when the signal-to-noise ratio is poor or with slow tissue motion [81]. The numerical and experimental images presented in the followings are reconstructed using a pixel size of 100 µm to maintain the ray tracing accuracy (discussed in Section III.C.).

The PC connected to the vantage has an Intel® Xeon® E5-2680 v3 CPU and a GeForce RTX-2080Ti GPU. As our MEX function is developed based on the direct sampling concept [77], the numerical dataset obtained by a sampling frequency of 533.3 MHz was down-sampled to a sampling frequency of 10 MHz, like in the experiments, and then passed to the MEX function.

### B. Evaluation metrics

The proposed method (ARC) aims to be much faster than our previous approach relying on an eikonal solver (FMT), while maintaining image quality. Therefore the difference between image quality metrics assessed in ARC images and in FMT images is reported. Lateral resolution calculated as the full-width-half-maximum (FWHM) and localization error (calculated with respect to the reference images) were used for quantitative evaluation of the image quality. The axial resolution is determined by the frequency bandwidth of the transmit waveform and is not significantly degraded by phase aberration. Therefore the axial resolution is not studied in this work.

To evaluate the effect of pixel size on the performance of the proposed ray tracing approach, the travel-times were calculated for a known velocity model (Figure 3(a)) using an Eikonal equation solver [82] and then compared to that obtained by our method. A spatial grid step size of 5 µm was used for the Eikonal solver. The model shown in Figure 3(b) is used to evaluate the effect of the pixel size on point spread function (PSF) and FWHM; ten point scatterers were positioned with lateral and axial coordinates of -3 mm to 5 mm and 20 mm to 36 mm, respectively.

The processing time of the MEX function was measured in full single element synthetic aperture imaging (SAI) and multi-angle plane (or diverging) wave imaging (PWI; with 11 virtual sources) schemes for depths of 20 mm up to 70 mm (suitable

<table>
<thead>
<tr>
<th>Scatterer number</th>
<th>Localization error (mm) [Lateral, Axial]</th>
<th>Lateral FWHM difference (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulation</td>
<td>Sawbones plate</td>
</tr>
<tr>
<td>1</td>
<td>[0.1, -0.1]</td>
<td>[0, 0]</td>
</tr>
<tr>
<td>2</td>
<td>[0, -0.4]</td>
<td>[0.2, 0]</td>
</tr>
<tr>
<td>3</td>
<td>[-0.1, -0.3]</td>
<td>[0, 0.1]</td>
</tr>
<tr>
<td>4</td>
<td>[-0.1, 0.3]</td>
<td>[0, 0]</td>
</tr>
<tr>
<td>5</td>
<td>[0.1, -0.2]</td>
<td>[0, -0.1]</td>
</tr>
<tr>
<td>6</td>
<td>---</td>
<td>[0, 0]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean ± standard deviation</th>
<th>Simulation</th>
<th>Sawbones plate</th>
<th>Skull</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0±0.1, 0.14±0.27]</td>
<td>[0.03±0.08, 0±0.06]</td>
<td>[1.2±0.31, 0±0.14]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean ± standard deviation</th>
<th>Simulation</th>
<th>Sawbones plate</th>
<th>Skull</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.04±0.05</td>
<td>0.07±0.03</td>
<td>0.05±0.14</td>
<td></td>
</tr>
</tbody>
</table>
for bone imaging [40] and transcranial imaging [83], respectively) with a pixel size of 100 µm and 200 µm. The medium configuration shown in Figure 3(b) was used. For each configuration, the MEX function was run 50 times and the average processing time was reported.

IV. RESULTS

A. Numerical results

1) Scatterers localization and FWHM

Figure 4(a, d) and Table 1 (the first column in the coordinates and FWHM sections) indicate that the proposed method does not significantly degrade image quality, as shown in Figure 1s (see the supplementary document). The mean lateral and axial localization errors are 0 mm and 0.14 mm, compared to the FMT, respectively. The mean difference in lateral resolution (FWHM) is 0.04 mm, compared to FMT.

2) Effects of pixel size

Figure 5(a) indicates that the pixel size does not significantly change the lateral resolution (FWHM). The PSF is not significantly altered (see Figure 5(b)). The maximum intensity of the PSF obtained with ARC is found very close to the true location of the scatterers indicated with the green triangle (from image without the aberrator).

Figure 6(c) shows the absolute error calculated by subtracting the travel times obtained by the proposed ray tracing technique (Figure 6(a)) and the Eikonal solver (Figure 6(b)); a pixel size of 200 µm was used for estimation of the travel-times. The symmetry of the error map is due to symmetry of the velocity model (see Figure 3(a)). For pixels close to width=0, refraction is very small such that the ray is nearly vertical.

Figure 5. (a) The lateral FWHM (see the colorbar for the scale) at different axial (depth) and lateral locations and for different pixel sizes, obtained with the numerical model shown in Figure 3(b); the black and red x-axes show the depth and the lateral locations of the scatterers positioned obliquely, respectively. (b) The point spread function (PSF) obtained for different pixel size; the scatterer was positioned at the lateral and axial distance of -3 mm and 20 mm, respectively. The green mark indicates the true location of the scatterer. The PSFs are shown with a dynamic range of 40 dB. PS stands for the pixel size.
because all layers are parallel to the array. Therefore, if the lateral coordinate of the pixel is equal to the lateral coordinate of the element, the error becomes extremely small; in this case, the error is minimum close to $x=0$ for a lateral pixel coordinate of -200 $\mu$m because the lateral coordinate of the array element is -150 $\mu$m.

Figure 6(d) shows the distribution of the errors obtained by different pixel sizes. The errors do not linearly increase with a larger pixel size since the boundaries of the velocity model at a depth of 3 mm and 6 mm are determined by the nearest pixel on the image grid (see Figure 3(a)); for a pixel size of 140 $\mu$m, it causes about 30 ns shift in the errors. Yet, the errors are below 45 ns, which is slightly larger than 10 % of the ultrasound wave period (10% of 400 ns for the P4-1 probe), which results in insignificant degradation of image quality [84].

B. Experimental results

1) Estimation of the aberrator thickness

Figure 7 compares the experimental reconstructed images using the two dataset (Sawbones plate and skull) and two reconstruction methods (ARC and FMT). A thickness of 4.2 mm and 1.46 mm was obtained for the Sawbones plate and the temporal bone of the skull, respectively (not shown here). The thickness of the Sawbones plate is in excellent agreement with that obtained with the FMT [38]. The skull thickness is about 0.16 mm larger with the ARC technique compared with the FMT [38]; this error is consider minor comparing to the pixel size of 100 $\mu$m. The image quality in Figure 7(a, c) is close to Figure 7(b, d), which is our goal. The red boxes show the effects of multiple scattering caused by the aberrator.

2) Scatterers localization and FWHM: Sawbones plate

Figure 4(b, e) and Table 1 (the second column in the coordinates and FWHM sections) indicate that the proposed method reconstructs the scatterers with an average lateral and axial localization error of 0.03 mm and 0 mm, compared to FMT, respectively. The mean difference in lateral resolution (FWHM) is 0.07 mm, compared to FMT.

3) Scatterers localization and FWHM: Human skull

Figure 4(c, f) and Table 1 (the third column in the coordinates and FWHM sections) indicate that the proposed method reconstructs the scatterers with an average lateral and axial localization error of -1.2 mm and 0 mm, compared to the FMT, respectively. The lateral error is 0.3 mm with respect to reference image (values are available in the supplementary document of reference [38]) though. The mean difference in lateral resolution (FWHM) is 0.05 mm, compared to FMT.

Figure 7. The experimental reconstructed image with the (a, b) Sawbones plate and (c,d) the real human skull in front of the probe. All the images are locally-normalized and log-compressed. The red boxes show the effects of multiple scattering caused by the aberrating layer (plate and bone). The numbers (in yellow) are referred for quantitative evaluation in Table 1 and Figure 4.
4) Power Doppler imaging

The reconstructed power Doppler images with or without the Sawbones plate in front of the probe are presented in Figure 8. The yellow arrows in Figure 8(a) show bubbles floating in the water tank. With the aberrator inserted, the clutter around the string is significantly weaker in the image reconstructed by the ARC technique, compared to the conventional reconstruction which ignores the aberrator. The width of string (measured as the FWHM in a direction normal to the string) was 0.37 mm with the ARC technique (Figure 8(c)), which is close to the width in the reference image (0.32 mm). The width of the wire obtained with conventional reconstruction appears wider (0.54 mm) as a consequence of degraded lateral resolution caused by phase aberration.

C. Processing time with the GPU implementation of the ARC technique

The processing time of the proposed reconstruction approach, calculated for the configuration shown in Figure 3(b), is presented in Figure 9. The red and black curves represent the processing time and frame rate, respectively. The frame rate with a pixel size of 100 µm is lower than with a pixel size of 200 µm, because the processing time increases with the number of pixels. Using SAI scheme with 96 single-element transmissions to reconstruct the image up to depths of 20 mm and 70 mm, a frame rate of 40 Hz and 32 Hz is obtained for a pixel size of 200 µm. The processing time for each step in the ARC technique for PWI mode and an imaging depth of 70 mm with a pixel size of 200 µm is presented in Table 2.

Table 2. The processing time of each step of the ARC technique for PWI with 11 transmit steered plane waves and an image depth of 70 mm

<table>
<thead>
<tr>
<th>Processing modules</th>
<th>Time [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary memory allocation</td>
<td>8</td>
</tr>
<tr>
<td>First layer ray-tracing</td>
<td>6</td>
</tr>
<tr>
<td>First layer reconstruction</td>
<td>1</td>
</tr>
<tr>
<td>Second layer ray-tracing</td>
<td>5</td>
</tr>
<tr>
<td>Second layer reconstruction</td>
<td>1</td>
</tr>
<tr>
<td>Third layer ray-tracing</td>
<td>6</td>
</tr>
<tr>
<td>Third layer reconstruction</td>
<td>2</td>
</tr>
<tr>
<td>memory allocation and deallocation, segmentation</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
</tr>
</tbody>
</table>

Figure 9. The processing time/ frame rate of the MEX function in (a) full single element synthetic aperture imaging and (b) plane wave imaging/ diverging waves imaging (with 11 steering angles) schemes for different imaging depths and pixel sizes (PS). The black and red graphs belong to the left (Processing time) and right (Frame rate) horizontal panels, respectively.
V. DISCUSSION

A. Imaging scheme

To calculate the round-trip travel-times that are required to reconstruct the image with a delay-and-sum algorithm, both the transmit and receive travel-times are needed. In full single element SAI scheme, the number of transmitters (Ni) and receivers (Nr) and their coordinates are the same. Therefore, the receive travel-times calculated with respect to Nr receivers (see the ray tracing kernel in Figure 2) can be used as transmit travel-times. In practice, however, single element SAI is likely not the optimal imaging strategy for TUI due to 1) a low signal-to-noise ratio (i.e., transmitting with a single element generates a low acoustic pressure in the brain) and 2) a low data acquisition frame rate, which does not allow blood flow quantification in the brain (usually used for diagnosis of brain disorders) [85].

To address these issues, PWI [86] and diverging waves imaging (DWI) [87] schemes have a great potential since 1) a good image quality can be achieved while a faster data acquisition allows capturing transient phenomena in the brain, as reported by Montaldo, et al. [88], and 2) transmitting with sub-apertures or with all the elements of the probe (with appropriate transmit delays) ensures a good signal-to-noise ratio. For TUI with a phased array transducer, DWI is more desired as it enables a larger field of view, compared to PWI. Using PWI/ DWI schemes requires estimating the transmit travel-times from virtual transmitters (usually defined behind the ultrasound probe) to pixels. This requires an additional ray tracing in steps 2, 4 and 6 of Section II.A. To do so, the ray tracing kernel in Figure 2 is configured with a Z grid size equal to the number of virtual transmitters.

As presented in Figure 9, PWI or DWI offers faster image reconstruction compared to single element SAI. That is because 1) the size of the third dimension (Z) of the reconstruction kernel (see Figure 2) equals to the number of virtual transmitter (from 3 [89] to 15 [90]), 2) less load/copy transactions from the global memory to on-chip memories, and the other way around, are needed. And, 3) the additional ray tracing (for estimating the transmit travel-times) is not computationally expensive. It should be noted that the processing time for PWI and DWI with the same number of virtual transmitters is the same.

B. Analytic signal calculation

Yiu, et al. [70] and Gonzalez and Bell [55] used Hilbert transform in the time and frequency domain to calculate the analytic echo data, respectively. In this study, however, the direct sampling concept is used to avoid dedicating processing power to calculate the analytic signal, which includes copy/load memory transactions as well [77]. The same methodology was used by Choe, et al. [62] before.

C. Limitations of the proposed approach

1) Wave-speed modeling

While the cortical bone of the temporal region of the skull is known to exhibit anisotropic wavespeed [73], the wavespeed in the bone layer was assumed isotropic to simplify ray-tracing and therefore enable the fastest frame rate with our approach. Image distortions (depending on the anisotropy level) are expected if the anisotropy is neglected.

2) Implementation

As can be seen in Table 2, there are two modules that increase the processing time: 1) primary memory allocation, 2) ray-tracing kernel.

a) Memory allocation

The primary memory allocation includes defining host (CPU) memories to save the results of the GPU kernels/ CPU functions and copying the properties of the imaging system to the GPU global memory (as shown in Figure 2). Currently, the memory allocation, and also memory freeing at the end of the MEX function, happens each time the MEX function runs. To further increase the frame rate, Matlab GPU arrays can be used to define these memories once (when the MEX function runs for the first time), and free them at the end of imaging. It should be also mentioned that the RF-data is transferred to the global memory in an asynchronous way during the first reconstruction step. Therefore, the size of the RF-data is not a bottleneck.

b) Ray-tracing kernel

Our ray-tracing kernel calculates the travel-time of all the pixels in parallel. However, finding the shortest travel-time between a pixel and an array element through IPs is implemented sequentially with a for-loop in GPU. This kernel can be improved by re-defining the kernel architecture in a way that it evaluates all the travel-times (through the IPs) in parallel. Another option is to implement it on CPU with parallel programming (such an OpenMP). In the current implementation, ray-tracing is calculated for each frame because the position and geometry of the skull (in the current image plane) must be updated to achieve real-time refraction-corrected TUI as the sonographer moves the probe to search for the optimal image plane. However, once the proper imaging window and imaging plane are found by the operator and the probe is stabilized on the temporal bone, the travel-times could be only calculated once and reused as long as the probe remains still. An automated approach could be implemented to evaluate the stability of the probe in order to decide whether or not the refraction-corrected travel times must be updated. This would further increase the frame rate of our method.

D. Future work

Our future studies include the in vivo evaluation of the proposed approach in adult human subjects, the development of a GPU-accelerated reconstruction technique that addresses the wavespeed anisotropy in the bone layer and the bottlenecks discussed in Section IV.C.2, and the translation of the approach to a matrix array transducer where the 3D geometry of the skull will be described.

VI. CONCLUSION

In this paper, we reported a reconstruction technique for transcranial ultrasound imaging through the temporal bone, which uses a fast ray-tracing approach and a GPU implementation to speed up the calculation of the refraction-corrected arrival-times and image reconstruction. The results showed that the error on travel-times using a pixel size of 200 μm is not significant (about 10% of the ultrasound temporal period at 2.5MHz), and hence resulting in insignificant
degradation of image quality and resolution, and negligible localization errors. We believe the approach enables real-time refraction-corrected transcranial ultrasound imaging with a frame rate between 20 and 30 Hz, and real-time refraction-corrected power Doppler imaging.

ACKNOWLEDGEMENT

This work was supported by a joint grant from the Netherlands Organisation for Scientific Research (NWO)/the Netherlands Organisation for Health Research and Development (ZonMw) and the Department of Biotechnology (Government of India) under the program Medical Devices for Affordable Health (MDAH) as Project Imaging Needles (Grant Number 116310008).

REFERENCES

[36] G. Bouchoux et al., "Experimental validation of a finite-difference model for the prediction of transcranial ultrasound fields based on


