Real-time Shear Wave Elastography Implementation on a Portable Research Ultrasound System with GPU-accelerated Processing

Damian Cacko[1, 2], Piotr Jarosik[1, 2], and Marcin Lewandowski[1, 2]
[1]*Institute of Fundamental Technological Research of Polish Academy of Sciences*, Warsaw, Poland
[2]*us4us Ltd.*, Warsaw, Poland
Email: damian.cacko@us4us.eu

***Abstract*—In this work, we present a low-cost, portable, and fully configurable ultrasound system implementing 2-D real-time Shear Wave Elastography (SWE) imaging mode. To achieve that we have enhanced the transmit capabilities of the 256 TX/64 RX us4R-lite research system, developed by our team, to support push pulses generation. This system was combined with a signal processing pipeline reconstructing stiffness maps from raw RF data. Real-time imaging performance was provided by an efficient reconstruction algorithm execution that incorporated graphics processing unit (GPU). The overall system performance was assessed experimentally using an industry-standard elasticity Q/A phantom. Relevant reconstruction parameters were evaluated in terms of reconstruction time. The system achieved stiffness estimation with a bias <5% and SNR of 30 dB and was able to detect lesions of size >4 mm and various stiffness with CNR in the range of 13–17 dB. The system throughput of up to 5 fps has been achieved on a PC notebook equipped with NVIDIA RTX 3060 GPU.**

***Index Terms*—Shear wave elastography, GPU, real-time imaging, ultrasound imaging systems design.**

I. Introduction

Shear wave elastography (SWE) quantifies tissue stiffness by measuring a shear wave speed (SWS) resulting from an acoustic radiation force (ARF) excitation. Dedicated methodologies have been developed to obtain 2-D maps of the underlying tissue stiffness, e.g. SuperSonic Shear Imaging (SSI) [1], or Comb-Push Ultrasound Shear Elastography (CUSE) [2]. The SWE found application in many fields, such as liver fibrosis staging [3], breast cancer detection [4], imaging of thyroid nodules [5], and many others [6].

The practical realization of a real-time SWE technique faces a number of technical challenges [7]. First, a system implementing this mode must have robust transmit capabilities that enable high-energy push-pulse generation for ARF, strong enough to induce detectable shear wave motion in a tissue. Then, one needs to provide the ability to capture channel data at a high frame rate (usually >=5000 fps) for effective shear wave tracking. In real-time operation, captured data must be transferred to the PC for further processing via a high-throughput interface. Finally, due to the huge amount of data to be processed to obtain a single tissue stiffness map, a very high computational performance of the data processing pipeline is required.

Due to the aforementioned reasons, up until now, the real-time 2-D SWE has only been accessible on few high-end specialized systems [8]. Those devices are clinical systems and their imaging implementation is hidden from the user, algorithm control or customization is not provided in any way, and the user has no access to raw channel data. Thus, their usefulness to effectively facilitate the development and evaluation of new diagnostic ultrasound methods and algorithms is very limited [9].

To address this, in this work, we developed the 2-D SWE imaging mode using a low-cost, portable, and fully configurable research system and utilized a graphics processing unit (GPU) to provide performance high enough for real-time imaging.

II. Materials and Methods

A. Hardware setup

A complete test setup used in this work is shown in Fig. 1. For data acquisition, we employed the us4R-lite system (us4us Ltd., Warsaw, Poland) developed by our team. This 256 transmit (TX) and 64 receive (RX) channels research platform has been optimized with respect to size, power, and cost. It features fully programmable TX/RX sequences, a high-speed data transfer to the PC via a 2.4 GB/s Thunderbolt-3 interface, and provides access to raw RF channel data, which allows implementation of any kind of custom processing in software. In addition, specifically for SWE, the system transmit capabilities have been enhanced, enabling the generation of high-energy push pulses. A summary of its transmit capabilities is presented in Table I.

An evaluation of the system performance was carried out using a 128-element ATL L7-4 linear array (Philips Healthcare, Amsterdam, The Netherlands) and a tissue-mimicking Model 049A elasticity Q/A phantom (CIRS, Norfolk, VA, USA). Stepped cylinder mass targets of four different diameters (4.0, 6.5, 10.4, and 16.7 mm) and known stiffness were used (70.9 kPa, or 4.86 m/s shear wave speed), all centered at a depth of 30 mm. The phantom region with no inclusions (background) of nominal stiffness of 16.7 kPa (2.36 m/s) was used in the homogeneous experiment.

The MSI GS66 mobile workstation (Micro-Star Int'l Co. Ltd., Taipei, Taiwan) was used as a host, equipped with an

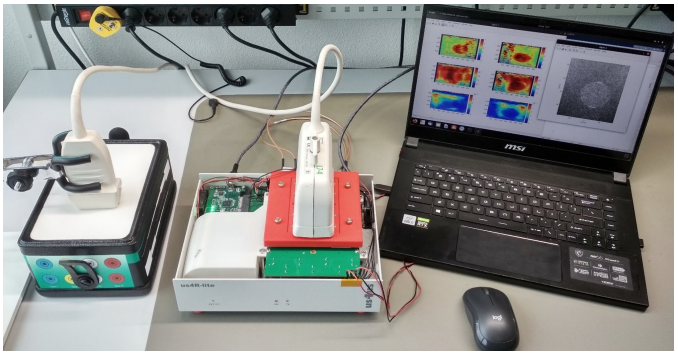

Fig. 1. The setup that was used in this work. Left to right: the ATL L7-4 probe over the CIRS 049A elastography phantom, the transmit-enhanced us4R-lite portable research platform connected to PC. Detailed description in the text.

TABLE I
ENHANCED US4R-LITE SYSTEM TRANSMIT CAPABILITIES

Parameter	Value
Excitation voltage	up to ± 160 Vpp
Output current	up to ± 4 A
Transmit aperture	1–256 elements
Transmit aperture apodization	Yes, PWM
Output frequency range	0.1–10 MHz (pushes)[a] 0.1–30 MHz (short cycles)[a]
Max. push duration	1 ms[a,b]
Max. tracking PRF	20 kHz[c]

[a]Depends on load. [b]Depends on push PRF.
[c]Depends on acquisition depth and receive aperture.

Intel Core i9-10980HK CPU with 18 cores, Nvidia RTX3060 GPU, and 32 GB of RAM memory.

B. Data acquisition

TX/RX sequences used to obtain a single stiffness map consisted of push generation and high frame rate tracking sequence. The push generation method followed the principle of the CUSE method enabling the reconstruction of the SWS map within the whole field of view (FOV) from a single push sequence. There were 2 parallel, laterally spaced pushing beams generated simultaneously, each using 64 transducer elements and focused at a depth of 35 mm, ($F/\# = 1.8$). Push duration was set to 350μs. A push pulse frequency of 4.4 MHz was used.

Immediately after the end of push generation, the us4R-lite scanner switched to high frame-rate plane wave imaging (PWI) mode transmitting 2 cycles at a center frequency of 4.4 MHz with a pulse repetition frequency (PRF) of 12 kHz using all probe elements and capturing echo from up to 50 mm depth. As proposed in [10], the coherent plane-wave compounding (CPWC) method was utilized with a set of angles $(-4^o, 0^o, 4^o)$. In each acquisition, a total number of 100 frames were captured. Due to a limited number of 64 RX channels of the us4R-lite system, each transmission was repeated twice to capture the whole 128-element receive aperture, which reduced the effective frame rate twice, down to 6 kHz, and required total of 200 plane-wave transmissions per acquisition.

C. Data processing system architecture

The developed processing back-end is presented in Fig. 2. Our implementation was a streaming processing system, where the acquired data were continuously transferred to the PC equipped with a GPU for image reconstruction. Reconstruction of a single stiffness map is a complex algorithm that involves the processing of tens of ultrasound frames. Thus, to reduce reconstruction time, the processing pipeline consisting of multiple stages, as shown in the image, was entirely run within the GPU. The GPU kernels were developed in Python environment and CuPy (Preferred Networks Inc., Tokyo, Japan) library was used to move the execution of computational-intensive calculations from CPU to GPU.

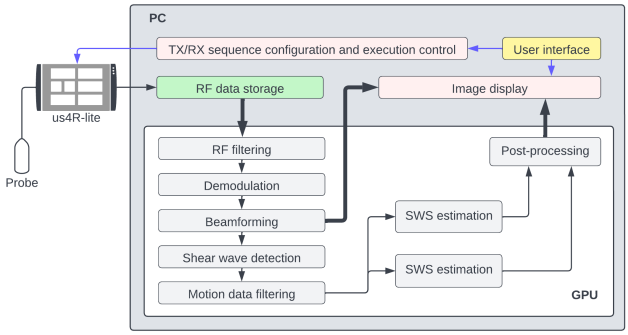

Fig. 2. Top-level architecture of the GPU-accelerated reconstruction pipeline. Blocks marked in light pink were executed solely on the CPU, while the gray blocks were implemented as kernels for execution on the GPU. For brevity, some intermediate processing steps were skipped in the diagram.

In the first step, the frames of raw RF channel data dumped by the us4R-lite system to the host memory after the acquisition were transferred to the GPU memory. Next, to improve the SNR of the input data, a symmetric pass-band 4–7 MHz 128-tap FIR filtering was applied to filter out frequency components coming from electronic noise. Quadrature demodulation was used to bring the signal to the baseband. In-phase and quadrature components of I/Q data were then filtered by a 128-tap low-pass FIR filter to cut off the demodulation products at the doubled fundamental frequency of the signal. These signal conditioning steps were followed by delay-and-sum CPWC beamforming. The reconstruction was done on the grid with a pixel size of 0.2 x 0.2 mm and grid size of 38 mm and 42 mm in lateral and axial directions, respectively. A walking average approach was utilized at the angle compounding of the beamformed data stage. Then, Kasai's 1D-autocorrelator algorithm in progressive referencing was utilized to obtain shear wave motion data from beamformed frames. Axial kernel size of 0.8 mm and ensemble length of 4 frames were used. To separate leftward and rightward propagating waves of the spatio-temporal motion data, the signal was transferred to the (k, ω) space using 2-D FFT and a directional filter similar to proposed in [11] was applied. In addition, the mask was

designed to remove all the signal components of frequencies below 40 Hz and above 800 Hz and to filter out all data components related to shear waves propagating speeds out of range of interest (0.5–4.5 m/s). As a result of directional filtering, two datasets were constructed — related to leftwards and rightwards propagating waves. Then, to obtain local SWS values a correlation-based time-of-flight estimation algorithm was implemented. A normalized cross-correlation of Tukey-windowed tissue motion profiles at laterally distant locations was calculated to find propagation delay at a known distance defined by lateral kernel size of 3.2 mm. In order to speed up the process, the correlations were found using an FFT-based approach. Before that, shear wave motion data was interpolated by a factor of 5 along the slow-time dimension using the spline interpolation method. The above algorithm was applied separately on two input datasets coming from the directional filtering step, producing for each an SWS map and a map of correlation coefficients that are treated as the estimation quality indicator maps. Multiple SWS maps were combined to build a single compounded SWS map as the weighted sum, taking the correlation coefficients as weights. Finally, in order to refine the image, median filtering with a kernel size of 1 mm x 1 mm was applied to the SWS map.

III. Results

The developed solution enabled to obtain SWS maps in the real-time regime, what was verified experimentally using an elastography phantom. Fig. 3 shows a snapshot from the display during ongoing imaging. Three images were displayed — a B-mode image for reference and probe positioning, a reconstructed SWS map, and an animation made of reconstructed shear wave motion data. With frame rates operation higher than 1 fps the animation was replaced on display by a selected shear wave motion data frame.

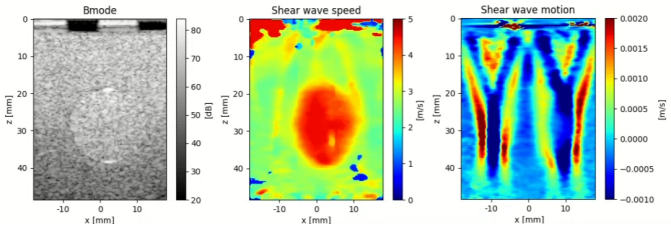

Fig. 3. Images displayed during real-time 2-D SWE imaging using the developed system. Left: B-mode image. Middle: SWS map of the 70.9 kPa inclusion of 16 mm diameter. Artifacts are visible. Right: Shear wave motion data frame or animation.

The system was able to refresh the display with a frame rate depending on user preference and values of reconstruction parameters. For the default set of parameter values, as described in section II-C, the total processing time required to reconstruct a single SWS map was 348 ms, which corresponds to 2.87 fps. The bar plot presented in Fig. 4 depicts how each of the reconstruction algorithm's steps contributes to the overall processing time. The SWS estimation was found the most computationally-intensive step that required almost 200 ms being this way responsible for more than half of the total processing time (57%).

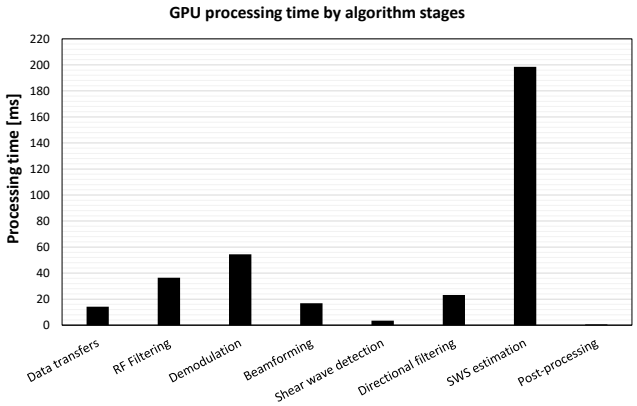

Fig. 4. Contribution of each reconstruction algorithm step to the total processing time for default values of reconstruction parameters. The data transfers bar includes both input and output data transfers.

Fig. 5 shows how the processing time of 2-D SWE-specific steps was influenced by the computational problem size. The size of the images to reconstruct was controlled by the beamforming grid step, which dictated the number of pixels to be reconstructed in all subsequent steps. Fig. 6 highlights how significantly interpolation parameters applied for the SWS estimation stage can affect the processing time; e.g., it requires 580 ms to execute when using order 3 spline interpolation with a factor of 10, whereas it needs only 200 ms to run with interpolation factor of 5 and spline interpolation order of 2, being almost 3 times faster.

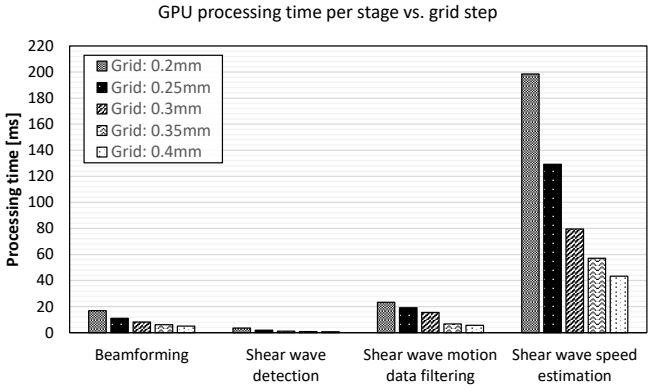

Fig. 5. Evaluation of computational problem size impact on processing time of selected algorithm steps. Dataset size was controlled by manipulating the beamforming grid step.

Regarding the imaging quality, the system achieved stiffness estimation with a bias <5% and SNR of 30 dB when imaging the homogeneous region of the phantom. In the case of imaging heterogeneous regions, it was able to detect lesions of size >4 mm and various stiffness with CNR in the range of 13–17 dB.

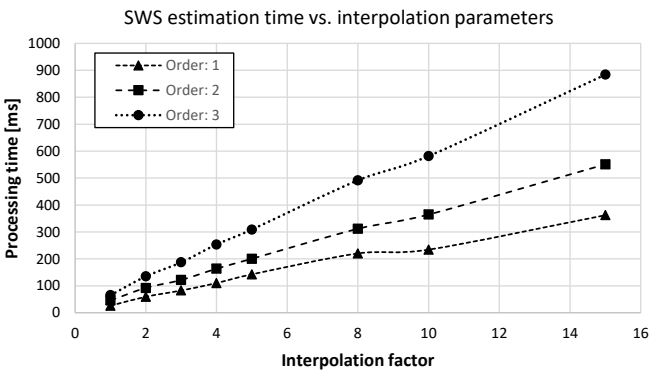

Fig. 6. SWS estimation processing time for various interpolation factors and the order of spline interpolation.

IV. Discussion

The presented system architecture benefits from the software-defined processing — CPU and GPU co-processing. The main advantage of this approach was shown in this study — a complex 2-D SWE imaging mode was developed using a high-level scripting language (Python), providing flexibility and scalable performance at the same time thanks to GPU usage. Our experiments revealed that GPU acceleration boosted the processing performance by a factor of 85x and 100x versus CPU-only Python and Matlab implementations of the same algorithm, respectively. The presented implementation took advantage of a high data parallelism inherent to the 2-D SWE image reconstruction algorithm, which allowed us to leverage GPU performance to speed up the reconstruction.

A significant impact of reconstruction parameters on data processing time was demonstrated in the presented results. Considering the computational performance only, the system proposed in this work can potentially provide 2-D elasticity maps with an image refresh rate of up to 10 fps, depending on reconstruction parameters settings. However, operating at such a high frame rate can result in poor imaging quality, excessive imaging system, and probe heating, and may cause exceeding regulatory acoustic output safety levels.

This study has several limitations. Firstly, imaging quality assessment was very limited. However, detailed results of system imaging quality assessment were published elsewhere ([7]). Secondly, there were experiments conducted to evaluate the system's performance including evaluation of key reconstruction parameters on processing time, but the impact of these parameters on imaging quality was not evaluated. This subject was beyond the scope of this work and research in this field is planned to be held in the future.

V. Conclusions

We developed a high-performance 2-D SWE imaging system using a low-cost portable research platform coupled with GPU processing on a PC notebook. Due to its high configurability, easy customization, and high performance, we believe, that this solution holds significant potential as a tool for the development of novel techniques and innovations in the field of SWE.

References

[1] J. Bercoff, M. Tanter, and M. Fink, "Supersonic shear imaging: a new technique for soft tissue elasticity mapping," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 51(4), pp. 396–409, 2004.

[2] P. Song, H. Zhao, A. Manduca, M. Urban, J. F. Greenleaf, and Shigao Chen, "Comb-push ultrasound shear elastography (CUSE): a novel method for two-dimensional shear elasticity imaging of soft tissues," IEEE Trans Med Imaging, vol. 31(9), pp. 1821–1832, 2012.

[3] R. G. Barr, "Shear wave liver elastography," Abdominal Radiology, vol. 43(4), pp. 800–807, 2018.

[4] R. G. Barr et al., "WFUMB Guidelines and recommendations for clinical use of ultrasound elastography: Part 2: breast," Ultrasound Med. Biol., vol. 41(5), pp. 1148–1160, 2015.

[5] D. Cosgrove et al., "WFUMB Guidelines and recommendations on the clinical use of ultrasound elastography: Part 4. thyroid," Ultrasound Med. Biol., vol. 43(1), pp. 4–26, 2017.

[6] D. Cosgrove et al., "EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography. Part 2: clinical applications," Ultraschall Med., vol. 34(3), pp. 238–245, 2013.

[7] D. Cacko and M. Lewandowski, "Shear wave elastography implementation on a portable research ultrasound system: initial results," Appl. Sci., vol. 12(6210), 2022.

[8] M. W. Urban, "Production of acoustic radiation force using ultrasound: methods and applications," Expert Rev. Med. Devices, vol. 15(11), pp. 819–834, 2018.

[9] E. Boni, A. C. H. Yu, S. Freear, J. A. Jensen, and P. Tortoli, "Ultrasound open platforms for next-generation imaging technique development," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 65(7), pp. 1078–1092, 2018.

[10] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 56(3), pp. 489–506, 2009.

[11] T. Deffieux, J-L Gennisson, J. Bercoff, and M. Tanter, "On the effects of reflected waves in transient shear wave elastography," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 58(10), pp. 2032–2035, 2011.