Single-FPGA Complete 3D and 2D Medical Ultrasound Imager

A. Ibrahim*, W. Simon*, D. Doy*, E. Pignat*, F. Angiolini*, M. Arditi*, J.-P. Thiran*†, and G. De Micheli*

* École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

†Department of Radiology, University Hospital Center (CHUV) and University of Lausanne (UNIL), Switzerland

***Abstract*—3D ultrasound (US) acquisition acquires volumetric images, thus alleviating a classical US imaging bottleneck that requires a highly-trained sonographer to operate the US probe. However, this opportunity has not been explored in practice, since 3D US machines are only suitable for hospital usage in terms of cost, size and power requirements. In this work we propose the first fully-digital, single-chip 3D US imager on FPGA. The proposed design is a complete processing pipeline that includes pre-processing, image reconstruction, and post-processing. It supports up to 1024 input channels, which matches or exceeds state of the art, in an unprecedented estimated power budget of 6.1 W. The imager exploits a highly scalable architecture which can be either downscaled for 2D imaging, or further upscaled on a larger FPGA. Our platform supports both real-time inputs over an optical cable, or test data feeds sent by a laptop running Matlab and custom tools over an Ethernet connection. Additionally, the design allows HDMI video output on a screen.**

I. Introduction

Telesonography, i.e. remote ultrasound (US) diagnostics, is a form of telemedicine with promising applications. Indeed, US imaging is fully safe, unlike ionizing radiation-based techniques like X-rays, and entails lower complexity and constraints during the scan than e.g. Magnetic Resonance Imaging (MRI). For these reasons, US imaging is highly prevalent in many medical fields. Telesonography would be extremely useful e.g. in emergency rescue operations (helicopters, ambulances) and rural areas. However, at the moment, a well-trained sonographer must finely manipulate the probe until the desired 2D body section is captured. This means that the presence of a trained expert is mandatory during the acquisition stage, and telesonography remains problematic.

3D US imaging, a technology originally developed mainly for cardiology applications, can solve this problem by acquiring volumetric images at once. The 3D nature of the scans relaxes the precision constraints on the positioning of the probe, potentially allowing untrained or lightly-trained operators, e.g. paramedics, to acquire scans on the field. Either in real-time or subsequently, these volumes could be uploaded to a hospital, where a resident sonographer would reach a diagnosis. Unfortunately, today's 3D US systems are only available in well-equipped hospitals, due to their very expensive, stationary and power-consuming embodiments. In turn, these properties are chiefly the result of the enormous computation requirements of volumetric US reconstruction.

US imaging is performed through three main steps: acquisition, reconstruction, and visualization. Acquisition happens via a process called *insonification*, whereby a 2-20 MHz sound wave is emitted by an array of piezoelectric elements (*probe*) into the Region Of Interest (ROI). This wave is reflected by body tissue inhomogeneities, that act as point sources (*scatterers*). The reflected echoes, still in the radio frequency (RF) range, are received by the same probe and the signals are then sent to the backend imaging system. Although techniques vary, the main processing step in any imager is *beamforming* (BF). BF is the process of mapping the echo signals to their origin (scatterer) by summing them according to a certain delay profile. The reflectivity of the scatterer correlates to the amplitude of the summed echoes. A key part of BF is thus the calculation of one delay profile per focal point. Another major part is *apodization*, i.e. a weighting applied to the echoes to suppress the side-lobes due to the limited directivity of the piezoelectric elements, while keeping the main-lobe as narrow and high as possible. Finally, a visualization processing step is performed to produce a displayable image. In particular, the image must undergo *scan conversion (SC)*, i.e. a polar-to-Cartesian transform, and *log-compression*, i.e. its brightness range must be mapped to a logarithmic scale adapted to human vision.

The most critical part of BF is delay calculation. The delay profiles along which echo samples must be summed represent the two-way time-of-flights of the US waves emitted from an origin O to a scatterer $S \in V$ and back to a transducer element $D \in 1, ...N$. The BF process can be expressed mathematically by (1) where $e(D, t_p)$ is the echo received by an element D at a time-of-flight t_p, and $w(S, D)$ is the weighting applied to the echoes (refer to Section III-D).

$$s(S) = \sum_{D=1}^{N} e(D, t_p(|\vec{OS}|) + t_p(|\vec{SD}|))w(S, D), \forall S \in V \quad (1)$$

In 2D imaging, to reconstruct a cross-section of 64×500 pixels at a rate of 50 frames per second (fps) using a 64-element probe, 102.4 M delays/s need to be calculated. On the other hand, for 3D imaging, to reconstruct a volume of $64 \times 64 \times 500$ voxels using a 32×32 matrix probe at a reconstruction rate of 50 frames per second (fps), we will need to calculate 104.8 trillion delays/s. In other words, from 2D to 3D imaging, as both the number of probe elements and the number of image points increase by about 1.5 orders of magnitude, the total number of calculations increases by a factor of 1000. This is a key reason for the cost and bulk of 3D US imagers.

Each delay value is calculated as the Euclidean distance from O to S then S to D (i.e. two square roots) divided by the speed of sound c in the medium, that is typically assumed constant. The $|\vec{OS}|$ distance must be computed much less frequently than $|\vec{SD}|$, being O constant, and can thus be disregarded. Still, this means that 3D imaging requires calculating more than 100 trillion $|\vec{SD}|$ square roots per second, which is extremely challenging to fit into a portable, cheap, and battery-powered device. The most common workaround used by existing medical systems is to pre-process (multiplex or accumulate) the transducer element signals, reducing their count from several thousands to few hundreds of channels. This keeps in check the computation requirements, albeit at an image quality cost. Different methods have been proposed

978-1-5386-3534-6/17/$31.00 © 2017 IEEE

to achieve this reduction, like analog pre-BF [1]–[4], multiplexing [5], and sparse 2D arrays [6].

In this work, we propose a complete, fully digital imaging system, capable of both 2D and 3D US reconstruction, that is able to process signals from up to 1024 transducer channels in a single Kintex UltraScale KU040 FPGA [7] with an estimated power consumption of around 6 W. The proposed platform supports real-time inputs via an optical connection, as well as offline simulation inputs over an Ethernet cable. The platform also supports video output to a High-Definition Multimedia Interface (HDMI) screen. The design is highly scalable for various probe dimensions (i.e. array or matrix probes for 2D or 3D imaging, respectively), and various numbers of transducer elements. To the best of our knowledge, this is the first imager capable of high-channel-count 3D reconstruction running on a single FPGA, and therefore meeting the cost and power constraints of a portable telesonography system.

The paper is organized as follows: a brief literature review of recent commercial and research work is presented in Section II. A full description of the proposed system including the optimization of each processing step and the corresponding FPGA architecture is shown in Section III. This is followed by the experimental results, in Section IV, and our conclusions.

II. Previous Work

To the best of our knowledge, all 3D US imaging systems, both commercial and research, handle the high input channel count (thousands of elements for a matrix array) by immediately reducing the number of channels to a few hundreds within the piezoelectric probe. This can be achieved with different techniques. Analog pre-BF [1]–[4] is the most common; it adds up the signals received by a group of transducer elements, according to a fixed delay profile, and connects the output to a single channel towards the imager. This reduces the wiring and computation complexity, but degrades the ability to focus and therefore impacts resolution. Sparse 2D-arrays [6] and multiplexing [5] similarly lower the number of channels, by choosing the count and pattern of active receive elements per scan. Again, the main drawback is the loss of information and the reduced image quality.

Despite this reduction of receive channel count, these systems still end up bulky and power-consuming. On the commercial side, [8]–[10] are very advanced and provide full 3D support, but they are aimed at hospitals: they come on a base with wheels, they are very power-consuming, and are very expensive. Many commercial US machines that are portable, like the GE Voluson i [11], Samsung UGEO HM70A [12], [13], and Chison Q9 [14], are essentially 2D imagers; they nominally offer a 3D feature by supporting mechanically-swept arrays with low channel count (e.g. 128 elements). This type of probes introduces motion artifacts, and is only suitable for applications like obstetrics, where the subject (a baby) is mostly still. Applications like cardiac imaging require a high frame rate and high resolution at the same time, and therefore full 3D support. For example, the Philips CX50 [15] has a mode for 3D Trans-Esophageal Echocardiography (TEE), leveraging a matrix probe of 2500 elements, but analog pre-BF is exploited to compress this data onto only 100 receive channels [16].

A few 3D US research systems have been proposed. They all reduce the receive channel count, yet still ending up bulky. For example, the recent second-release ULA-OP system [17] is able to perform 256-channel BF on 8 high-end FPGAs and 16 DSPs. The advanced research platform SARUS [18] supports 1024 receive channels, the highest count supported by any 3D system, but requires 320 FPGAs. The Sonic Millip3De system [19], [20], that performs ultra-fast imaging, uses 128×96 probe elements - but only 1024 channels are considered per shot - with a powerful die-stacked package. However, its main bottleneck is the required external DRAM memory to store the BF delay coefficients, with the need of several GB/s memory bandwidth. There are also few single-FPGA US research systems, but they support up to only 64 receive channels. Moreover, there are many US systems relying on software-based BF, running on GPU, CPU, or DSP, but the resulting power budget is not optimal for battery operation, specially in 3D imaging.

A smart and efficient BF technique is needed to achieve a compact design, yet high-quality reconstruction. Each processing step in the imager should be optimized to yield overall platform compactness (ideally, single-chip) while being compatible with battery operation. In this work, we propose a novel, fully-scalable, complete, and single-FPGA US back-end system. We built our system around our previously proposed BF technique [21]–[23]. In this design, an efficient and optimized method for each step in the back-end system has been used to achieve overall platform efficiency, portability, and low power consumption.

III. Proposed US Imager System Architecture

In order to achieve a single-chip US processing platform that is able to support a high number of receive channels, in particular for 3D imaging, an optimization for each block in the back-end pipeline needs to be performed. This should be achieved while preserving image quality, since we target medical applications. We managed to fit the US processing logic for up to 1024 channels into a single latest-generation but mid-range FPGA, a Kintex UltraScale KU040 [7]. Figure 1 shows the architectural diagram of the proposed design.

A. System Specifications

Our US imager architecture is highly scalable and adaptable for different settings and parameters, e.g. the number of receive channels, the center frequency and sampling frequency, etc. However, in the following we will report numbers based on the settings specified in TABLE I. We present two versions of our design: a 2D US imager that supports 64 receive channels and reconstructs 64×500 focal points per frame, and a 3D US imager that supports 1024 receive channels and $64 \times 64 \times 500$ focal points.

B. Optical Cable Support For Probe Connection

In order to provide realtime data to the processing platform, an efficient communication channel with a probe should be chosen. Due to the high bandwidth requirements, especially for 3D imaging, the two preferred options to connect a probe to our KCU105 board [25] are either over the optical SFP ports or the Peripheral Component Interconnect Express (PCIe)

TABLE I
System Specifications

Parameter	Value
Speed of sound in tissue	1540 m/s
Transducer center frequency	4 MHz
Transducer bandwidth	4 MHz
Transducer array size (2D)	64 elements
Transducer matrix size (3D)	32×32 elements
Wavelength	0.385 mm
Sampling frequency	20 MHz
Focal points (2D)	64×500 = 32k
Imaging cross-section ($\theta \times r$, 2D)	73°×260λ
Focal points (3D)	$64 \times 64 \times 500$ = 2 M
Imaging volume ($\theta \times \phi \times r$, 3D)	73°×73°×260λ

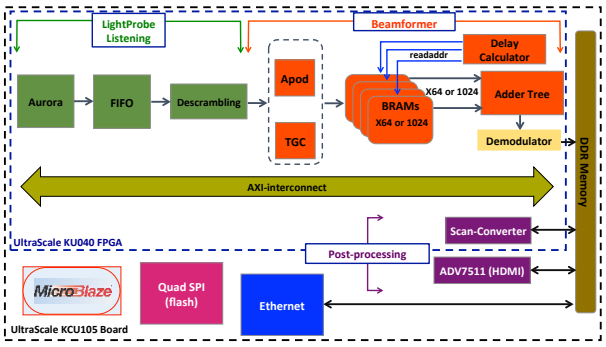


Fig. 1. Block diagram of the proposed US Imager.

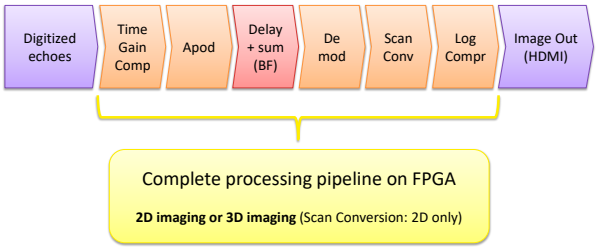


Fig. 2. Logical processing flowchart of the US image reconstruction pipeline.

interface. The latter implies higher complexity, in terms of design and cabling, as well as higher power consumption, so the optical interface is favoured.

Thanks to a collaboration with the Integrated Systems Laboratory (IIS) of the Swiss Federal Institute of Technology in Zurich (ETHZ), we have access to a 1D probe of 64 elements and 4 MHz center frequency with optical connection support [24]. The ADCs in the probe sample the data with 12-bit resolution at 20 MHz. The sustained output bandwidth is thus of around 15.3 Gbps. To support it, the Quad enhanced Small Form-factor Pluggable (QSFP+) interface is used by the probe; each of its four SFP+ lanes has a nominal bandwidth of 6.25 Gbps and a net one of 5 Gbps, for a total of 20 Gbps. Each lane is driven by a Xilinx Aurora 8b10b LogiCORE IP [26] for lightweight, scalable, and high-speed serial communication.

Within our Kintex FPGA, we have also implemented a 4-lane Aurora input channel. Since the KCU105 board [25] natively only features two $\times$ SFP+ ports, we have connected

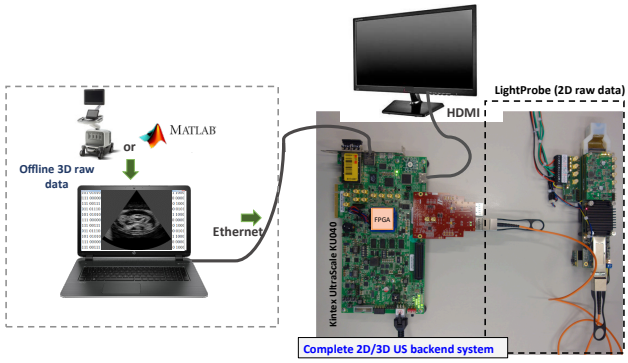


Fig. 3. Setup of the design for 2D and 3D US imaging. The components in the dashed box to the left are for development, debugging and verification only. For the probe on the right, which is not covered in this paper, refer to [24].

an intermediate QSFP+ board to the FPGA Mezzanine Card (FMC) interface [27] on the KCU105 (Fig. 3).

C. Ethernet Support For Simulated 3D Data Input

To the best of our knowledge, no openly accessible matrix probe for 3D imaging exists. To feed our imager with 3D data for development and debug, offline simulated data have been used. The data were reconstructed with the Field-II simulator [28] for Matlab to simulate matrix probes and 3D phantoms. For this 3D development mode, we have utilized an Ethernet port as the data transmission mean between a laptop and the FPGA (Fig. 3). Moreover, we have developed a C#-based Graphical User Interface (GUI). The GUI gives the user the control to choose among a small database of phantoms, pick imaging parameters, establish the Ethernet connection, and start sending the simulated raw data corresponding to the chosen phantom to the board. In addition, it allows the choice between different imaging modes (Section III-I) for both 3D imaging (i.e. simulated data) and 2D imaging (i.e. the 1D probe real data). The availability of simulated data is very beneficial as it provides a golden reference image, which helps in both debugging and quality assessment.

D. Static Apodization and Time-Gain Compensation

Both apodization and time-gain compensation are two weighting functions applied to the returned echoes (Fig. 1, Fig. 2). They can be expressed through the weighting parameter $w(S, D)$ in (1) of the BF process, as follows:

$$w(S,D) = w_{apod}(z_S, D) w_{tgc}(z_S), \forall D \in 1, ...N, S \in V, \\ and\ S = (x_S, y_S, z_S) \quad (2)$$

The apodization [29] w_{apod} is a weighting window - typically a Hanning function or similar - applied to the echoes received by each piezoelectric element D of the transducer. This compensates the antenna-like behavior of the element array, that yields side-lobes in the receive directivity. Typically, the apodization window has a width that is a function of the imaging depth (z_S), called "dynamic" or "expanding-aperture" apodization [29]. However, at a shallow depth - which is 1.6 cm with our settings - the window already reaches its maximum width, equal to the whole transducer's width. Thereafter, it becomes a static function of only D, $w_{apod}(D)$. In order to reduce the resources utilized by our imager, and since the difference between static and dynamic apodization is in the very shallow and less critical region for diagnosis, we chose to implement a fully static apodization in our imager. The static apodization coefficients, one per transducer element (therefore 64 in 2D and 1024 in 3D), can be pre-calculated and stored in a small table with 16-bit representation. The total table size is 128 B in 2D and 2 kB in 3D. This fits in a single Xilinx BRAM, whose size is 1024 rows of 18-bits. In order to apply the apodization coefficients to the RF data on the fly, a number of DSP multipliers matching the input data rate, i.e. one sample and one multiplication per clock cycle, must be deployed.

The US wave propagating inside the body is exposed to attenuation directly proportional to the traveling depth. This means that echoes returning from further depths (i.e. arriving later) are exposed to more attenuation than echoes returning from shallower depths (i.e. arriving earlier). A time-based compensation needs to be applied, called Time-Gain Compensation (TGC). TGC can be performed in the analog domain, for example in the transducer head, but also in the digital domain, as we implemented ($w_{tgc}(z_S)$). In the proposed

design, the TGC coefficients are also pre-calculated and stored in a small table; a single DSP multiplier has been used to apply these coefficients to the apodized RF data.

E. Steering-based Delay Calculation Algorithm

The main bottleneck and the most challenging process in the 3D US processing pipeline is delay calculation ($t_p(|\vec{OS}|)$ and $t_p(|\vec{SD}|)$). These delays are used as indices for the raw data to be accordingly summed (Eq. (1)) to determine both the scatterers' location and reflectivity. Delay calculation is the calculation of both the transmit Tx and receive Rx delays. For a specific emission origin O and to reconstruct a single volume V, Tx calculation is less demanding than Rx calculation by a factor of the number of transducer elements D. In 3D imaging and according to Table I, for a single transmission origin O, 2 million Tx delays need to be calculated versus 2 billion Rx delays *per volume*. Tx delays can be calculated on-the-fly, but for Rx calculation, at a target reconstruction rate of e.g. 50 vps, a computation bottleneck appears.

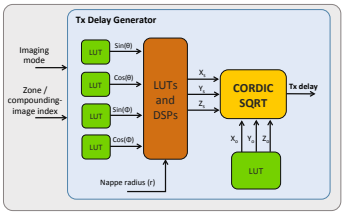
In the proposed design, the apodized and time-gain-compensated RF data are stored in BRAMs (Fig. 1), to be fetched by the delay calculator. In 2D imaging, we utilize one BRAM per receive channel, for a total of 64 BRAMs, while in 3D, to reduce BRAM pressure, we map every two receive channels to a single BRAM, i.e. using 512 BRAMs for 1024 channels. The Tx delays are calculated using directly a Xilinx CORDIC core, as seen in Figure 4(a). The location of each scatterer $S = (x_S, y_S, z_S)$ (as a function of the current azimuth and elevation angles θ, ϕ) and of the current emission origin $O = (x_O, y_O, z_O)$ are resolved and used as inputs. Depending on the imaging mode (Section III-I), the initialization of the azimuth θ and elevation ϕ pointers, as well as the emission origin, changes.

To solve the challenge of Rx delay calculation, we employ an efficient approximated algorithm [21]–[23]. It is based on the first order Taylor expansion of the square root. We simplify the calculation of the enormous number of square roots per second to the exact calculation of very few square roots along the central line-of-sight ($t_p(|\vec{RD}|)$), which is also done with a Xilinx CORDIC core, and then add two correction coefficients ($c_1 = \frac{x_D \sin\theta}{c}$ and $c_2 = \frac{y_D \sin\phi\cos\theta}{c}$) to calculate the delays for the remaining lines-of-sight (Eq. (3)). This can be seen as "steering" the delays of the central line-of-sight, where points R are located, to reconstruct the whole frame. Figure 4(b) shows the implementation of the Rx delay calculator. In (3), θ and ϕ are the azimuth and elevation angles, respectively, while x_D and y_D are the positions of each element D on the transducer face. r is the radial depth of the focal point being reconstructed, and c is the speed of sound in the medium, which is typically assumed constant. The architecture work identically for 2D imaging, setting ϕ and y_D to zero. Since θ, ϕ, x_D, y_D, r, and c have limited and deterministic values, the two correction coefficients can not only be calculated on-the-fly but also pre-calculated and stored in a small memory size, which we choose to do. The reconstruction rate of our beamformer is one voxel (or pixel) per clock cycle.

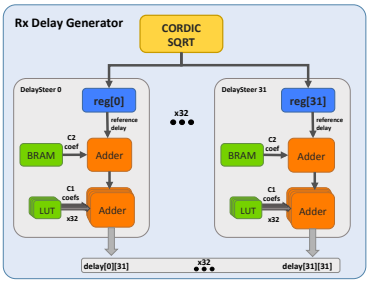
$$t_p(|\vec{SD}|) = r\sqrt{1 + \frac{x_D^2 + y_D^2}{r^2} - \frac{2x_D \sin\theta + 2y_D \sin\phi\cos\theta}{r}}$$
$$\approx t_p(|\vec{RD}|) - \frac{x_D \sin\theta}{c} - \frac{y_D \sin\phi\cos\theta}{c} \quad (3)$$

F. Demodulation Method

After BF, the reconstructed image is still in the RF domain, which needs demodulation, or in other words envelope detection. Many techniques are available. A very simple demodulation technique has been implemented to reduce the resource utilization and circuit complexity. It is based on simply calculating the absolute value of the reconstructed focal points, then applying a P-order FIR low-pass filter. After each clock cycle, the absolute value of the beamformed focal point is calculated and then stored in a circular buffer of depth $P+1$. The buffer is sufficiently wide to store a whole nappe of the volume, i.e. a surface with constant depth from the origin [30]. The P-order FIR filter is applied along the lines-of-sight of the buffered nappes. Finally, the demodulated focal points are then stored in another buffer until the completion of the whole volume/image, to be then used by the visualization unit (Section III-G).



(a)



(b)

Fig. 4. The steering-based delay calculation architecture. (a) Tx delay calculator, which is based on Xilinx CORDIC IP. (b) Rx delays are computed by adding the two steering coefficients c_1 and c_2 to exactly-calculated reference delays. The sum of Tx and Rx delays is used to index the input samples in the BRAMs of Fig. 1.

G. Cross-Sectional Scan-Conversion Block and HDMI Support

The last processing step in our US imager pipeline is visualization (Fig. 2). This includes two main processes; log-compression and scan-conversion (SC). The brightness of the US images just after demodulation has a very high dynamic range, which makes them appear either completely black or completely white to the human eye. The log-compression operation maps the image onto a logarithmic brightness scale, with appropriate contrast.

Another key operation is necessary since, in most US imaging methods including 3D US, the beamformed images are reconstructed in polar coordinates. SC transforms the image into the Cartesian coordinate space, to be displayable on a screen. This transformation is performed using interpolation, which also allows image scaling.

The SC process for a whole volume (i.e. 3D reconstruction) is computationally and materially expensive to be fit in a single FPGA. Luckily, this work aims at decoupling image acquisition and diagnosis. This means that the full 3D SC can be performed remotely at the hospital, when needed; the local operator needs at most a 2D cross-sectional image display for guidance and feedback. Therefore, we implement a block that

performs this operation. In the simpler case of 2D imaging, this block scan-converts the whole image.

The architecture of our SC block allows the operator to choose, using the on-board push-buttons, which cross-section of the volume is to be scan-converted and displayed. The default displayed cross-section by the system is the middle elevation slice (i.e. middle azimuth-depth plane) of the volume. Further, the proposed design supports HDMI output to a screen. We have used the Analog Devices ADV7511 [31] part on our KCU105 board for transmitting the HDMI output. The design also allows changing between different output resolutions from 640×480 to Full-HD 1920×1080.

H. System Self-Bootability

The FPGA imager is self-bootable, via a Quad Serial Peripheral Interface (QSPI) module that loads the boot software from the on-board flash memory.

I. Different Imaging Modes Support

Our single-chip architecture supports three different US imaging modes: single-insonification reconstruction, zone imaging, and compound imaging. They are considered among the essential modes in any US imaging system.

1) Single-Insonification Reconstruction: In this mode, the reconstruction of the whole volume/slice is performed based on a single insonification. To cover the whole ROI, a broad emission profile is used, such as a diverging or plane wave. This approach offers the highest possible reconstruction rate and has very low memory requirements. This comes at the cost of image quality, in particular very low lateral resolution, since the emitted acoustic energy is spread too broadly. This reconstruction technique is at the basis of ultrafast imaging [32].

2) Zone Imaging: In this mode, the ROI is divided into a number of non-overlapping zones, which are insonified in sequence by a focused beam, narrower than the one used in single-insonification. This improves the reconstruction quality, in particular providing enhanced lateral resolution. The acquisition time however becomes longer, proportionally to the number of zones and insonifications, while the processing time, processing resources, and memory requirements are almost the same. There are only two main differences in the processing pipeline: the first is the usage of different sub-Tx delay tables for each zone according to the different emission origins O, while the second is an extra "stitching" step, performed just after BF, to join the zones in memory.

3) Spatial Compound Imaging: Compound imaging performs BF for the whole image multiple times, but each time a different insonification profile is used. We have implemented spatial compounding based on steering the emission profile by 20° five times (i.e. five images to be compounded). We have used *averaging* as the compounding operator for the different images. Compound imaging has the advantage of reducing the speckle and clutter effects [33] of the US images, and it also improves the organ delineations and boundaries. Design-wise, compound imaging requires more acquisition time, processing time and/or resources, and memory requirements. This is because of the need for multiple insonifications, multiple full BF passes (except the SC step), and more storage requirements before the compounding step of the contributing reconstructions.

IV. Experimental Results

A. FPGA Architectural Results

We have evaluated the FPGA implementation results for the two versions of the proposed platform: a 64-channel 2D US imager, and a 1024-channel 3D US imager. We have used the Kintex UltraScale KCU105 as the deployment board, which is very well-suited to our architectural requirements. Table II shows the utilization of the two proposed imagers. The two most critical resources are the LUTs and the BRAMs. LUTs are mostly utilized by the beamformer and the AXI interconnect, which is configured with a "maximize performance" setting in Vivado, needed to support the high resolutions of the HDMI. As for BRAMs, 66% of the consumption is by the beamformer, with 42.7% (i.e. 512 BRAMs) just to store the received echoes, which is not further compressible. The architecture can be easily upscaled but may need a larger FPGA with more BRAMs and LUTs. In the third row of Table II, we have also extrapolated the results of upscaling the proposed imager to support 90×90 channels on a Virtex UltraScale XCVU190 FPGA.

In a real imaging system, the acquisition rate (i.e. the cadence of the sequence of beamformed images to be stored and uploaded to the hospital) and the display rate (i.e. the images to be displayed on a local screen, after SC, for the operator to monitor) can be decoupled. The proposed BF imager block runs at 133 MHz, and the BF rate is 1 pixel (or voxel) per clock cycle. This means that according to our system specifications, the *theoretical* reconstruction rate is 4157 fps in the 2D case, and 64.9 volumes/sec (vps) in the 3D case, which is more than sufficient to upload high-temporal-resolution streams of images (movies) to a hospital.

Looking at the complete system, when *inputs and outputs* are transmitted over the Ethernet port, which is a debug setup, this interface (and the related Microblaze software stack for packet processing) becomes a crucial bottleneck, resulting in a reconstruction rate of about 0.3 fps for 2D and 0.02 fps for 3D. When adding SC and using the HDMI port for direct image *output*, the reconstruction rate improves to 0.7 fps in either 2D or 3D, bottlenecked mostly by the Ethernet *inputs* and partially by the SC stage. Further optimizations to the SC block are ongoing. Our aim is to reach a 10 fps display rate in a realistic demonstrator featuring optical inputs and HDMI outputs.

The estimated power consumption - by Vivado - of our platform is 4.6 W for 2D imaging case and 6.1 W for 3D imaging case, which fully confirms the possible utilization in a battery-powered setup. As future work, we plan to directly measure the actual system power consumption of the board with a dedicated Maxim power measurement tool [34]. This will also require a careful optimization and power management of any board components and interfaces which are not needed by the US imager.

B. Quality Assessment

Our design achieves excellent resource efficiency at the cost of a slight inaccuracy in image reconstruction. The two main causes are the steering-based approximate delay calculation algorithm, and the static apodization. For the former, we have mathematically analyzed the image quality on reference images, and have confirmed [23] - as theoretically derived -

TABLE II
Imager Resource Utilization.
**Kintex UltraScale KU040 implementation results.*
***Virtex UltraScale XCVU190 extrapolated results.*

Supported Channels	Logic LUTs	Regs	BRAM	DSP	Clock	Theoretical Rate
64*	49.7%	22.3%	49%	7.2%	133 MHz	4157 fps
32×32*	85%	34.5%	97.1%	7.2%	133 MHz	65 vps
90×90**	84.3%	31.5%	89.4%	7.7%	133 MHz	65 vps

that the introduced inaccuracy occurs only very close to the probe surface and at the extreme lateral edges of the ROI. These regions are usually not critical for diagnosis, and in practice the disturbance is limited to a minor speckle pattern difference. Nonetheless, a detailed inaccuracy quantification using different metrics has been performed in [23], [35]. For what concerns static apodization, we have evaluated the impact of this optimization, finding that it only affects the shallowest 1.6cm of the ROI, which overlaps with the inaccuracy introduced by the steering-based algorithm. The same considerations apply.

V. Conclusions and Future Work

In this work, we have developed the first single-FPGA 1024-channel US imager. The proposed imager is able to perform real-time 2D and 3D complete US reconstruction including pre-processing, BF, and post-processing. The platform supports up to 1024 receive channels, which matches the state of the art, with an unprecedented estimated power consumption of 6.1 W. This has been successfully accomplished by utilizing an efficient delay calculation algorithm, in addition to optimizing each step in the processing pipeline. The design supports both real-time inputs, over an optical connection which we could test in combination with a probe for 2D imaging, and simulated inputs, thanks to the Ethernet support for both 3D and 2D imaging. The proposed platform enables telesonography by allowing the exploitation of ultrasound diagnosis by any untrained operator in remote rural areas, underdeveloped regions, rescue scenarios, and battlefields.

We plan to further optimize the platform in several respects, chiefly in terms of resource occupation. We also plan to more accurately measure and optimize the power consumption and to accurately assess the maximum achievable frame rate.

Acknowledgments

The authors acknowledge Swiss Confederation funding through the UltrasoundToGo project of the Nano-Tera.ch initiative.

References

[1] J. Larson, "2-d phased array ultrasound imaging system with distributed phasing," Jul. 20 1993, uS Patent 5,229,933. [Online]. Available: https://www.google.ch/patents/US5229933

[2] B. Savord and R. Solomon, "Fully sampled matrix transducer for real time 3D ultrasonic imaging," in *Ultrasonics, 2003 IEEE Symposium on*, vol. 1. IEEE, 2003, pp. 945–953.

[3] J.-Y. Um, Y.-J. Kim, S.-E. Cho, M.-K. Chae, B. Kim, J.-Y. Sim, and H.-J. Park, "A single-chip 32-channel analog beamformer with 4-ns delay resolution and 768-ns maximum delay range for ultrasound medical imaging with a linear array transducer," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 1, pp. 138 – 151, February 2015.

[4] H. gil Kang, S. Bae, P. Kim, J. Park, G. Lee, W. Jung, M. Park, K. Kim, W. Lee, and T.-K. Song, "Column-based micro-beamformer for improved 2D beamforming using a matrix array transducer," in *Proceedings of the 2015 Biomedical Circuits And Systems (BioCAS) Conference*, 2015.

[5] T. M. Carpenter, M. W. Rashid, M. Ghovanloo, D. Cowell, S. Freear, and F. L. Degertekin, "Time-division multiplexing for cable reduction in ultrasound imaging catheters," in *Proceedings of the 2015 Biomedical Circuits And Systems (BioCAS) Conference*, 2015.

[6] A. Austeng and S. Holm, "Sparse 2-d arrays for 3-d phased array imaging - design methods," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control.*, vol. 49, no. 8, pp. 1073–1086, Aug 2002.

[7] Xilinx Inc., "Ultrascale FPGA: Product tables and product selection guide," 2016, http://www.xilinx.com/support/documentation/selection-guides/ultrascale-fpga-product-selection-guide.pdf#KU.

[8] "EPIQ 7 ultrasound system," Philips Healthcare, 2017. [Online]. Available: http://www.medical.philips.com/main/products/ultrasound/systems/epiq7/

[9] GE Healthcare, "Voluson E10," 2017, www.gehealthcare.com.

[10] Samsung Electronics Co., Ltd., "Samsung WS80A with elite," 2017, http://www.samsung.com/global/business/healthcare/healthcare/ultrasound/USS-WS8AF4K/WR.

[11] GE Healthcare, "Voluson i," 2017, http://www3.gehealthcare.co.uk/en-gb/products/categories/ultrasound/voluson/voluson_i.

[12] Samsung Healthcare, "HM70A with plus," 2017, http://www.samsung.com/global/business/healthcare/healthcare/ultrasound/USS-HM7AF1K/WR.

[13] kpi Healthcare, "Samsung UGEO HM70A ultrasound machine," 2017, http://www.kpiultrasound.com/medical-equipment/ultrasound-scanners/samsung-ultrasound/samsung-ugeo-hm70a-portable/.

[14] Chison Medical Imaging Co., Ltd., "Chison q9," 2014, http://www.chison.com/products_detail.asp?id=168.

[15] Philips Healthcare, "Cx50 xMATRIX," 2017, http://www.usa.philips.com/healthcare/product/HC795076CA/cx50-xmatrix-cardiology-ultrasound-machine.

[16] G. Wayne Moore, "Ultrasound beamformation and image reconstruction," in the 51st American Association of Physicists in Medicine Annual Meeting (AAPM 2009), 2009, https://vimeo.com/88175266.

[17] E. Boni, L. Bassi, A. Dallai, F. Guidi, V. Meacci, A. Ramalli, S. Ricci, and P. Tortoli, "ULA-OP 256: A 256-channel open scanner for development and real-time implementation of new ultrasound methods," *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, in press 2016.

[18] J. Jensen, H. Holten-Lund, R. Nilsson, M. Hansen, U. Larsen, R. Domsten, B. Tomov, M. Stuart, S. Nikolov, M. Pihl, Y. Du, J. Rasmussen, and M. Rasmussen, "Sarus: A synthetic aperture real-time ultrasound system," *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, vol. 60, no. 9, pp. 1838–1852, Sep 2013.

[19] R. Sampson, M. Yang, S. Wei, C. Chakrabarti, and T. Wenisch, "Sonic millip3de: A massively parallel 3d-stacked accelerator for 3d ultrasound," in *High Performance Computer Architecture (HPCA2013), 2013 IEEE 19th International Symposium on*, Feb 2013, pp. 318–329.

[20] R. Sampson, M. Yang, S. Wei, R. Jintamethasawat, B. Fowlkes, O. Kripfgans, C. Chakrabarti, and T. Wenisch, "FPGA implementation of low-power 3D ultrasound beamformer," in *2015 IEEE International Ultrasonics Symposium (IUS 2015)*, Nov 2015.

[21] A. Ibrahim, P. A. Hager, A. Bartolini, F. Angiolini, M. Arditi, L. Benini, and G. De Micheli, "Tackling the bottleneck of delay tables in 3D ultrasound imaging," in *Proceedings of the 2015 Design Automation and Test in Europe (DATE 2015) Conference*, March 2015, pp. 1683 – 1688.

[22] F. Angiolini, A. Ibrahim, W. Simon, A. C. Yüzügüler, M. Arditi, J.-P. Thiran, and G. De Micheli, "1024-Channel 3D ultrasound digital beamformer in a single 5W FPGA," in *Proceedings of the 2017 Design Automation and Test in Europe (DATE 2017) Conference*, March 2017.

[23] A. Ibrahim, P. A. Hager, A. Bartolini, F. Angiolini, M. Arditi, J.-P. Thiran, L. Benini, and G. De Micheli, "Efficient sample delay calculation for 2d and 3d ultrasound imaging," *to appear in IEEE Transactions on Biomedical Circuits and Systems (TBCAS)*, 2017.

[24] P. A. Hager, C. Risser, P.-K. Weber, and L. Benini, "LightProbe: a 64-channel programmable ultrasound transducer head with an integrated front-end and a 26.4 Gb/s optical link," in *Proceedings of the 2017 IEEE International Symposium on Circuits and Systems (ISCAS 2017)*, 2017.

[25] Xilinx Inc., "Xilinx kintex ultraScale FPGA KCU105 evaluation kit," 2017, https://www.xilinx.com/products/boards-and-kits/kcu105.html.

[26] ——, "Aurora 8B/10B," 2017, https://www.xilinx.com/products/intellectual-property/aurora8b10b.html.

[27] HiTech Global LLC, "2-port QSFP+ (40G or 56G) FMC module (vita57.1)," http://www.hitechglobal.com/FMCModules/FMC_2QSFP+.htm.

[28] J. Jensen, "Field: A program for simulating ultrasound systems," in *10th Nordic-Baltic Conference on Biomedical Imaging Published in Medical and Biological Engineering and Computing*, vol. 34, Supplement 1, 1996, pp. 351–353.

[29] C. Daft and W. Engeler, "Windowing of wide-band ultrasound transducers," in *Ultrasonics Symposium, 1996. Proceedings., 1996 IEEE*, vol. 2. IEEE, 1996, pp. 1541–1544.

[30] P. Vogel, A. Bartolini, and L. Benini, "Efficient parallel beamforming for 3D ultrasound imaging," in *Proceedings of the 24th Edition of the Great Lakes Symposium on VLSI*, ser. GLSVLSI '14. New York, NY, USA: ACM, 2014, pp. 175–180. [Online]. Available: http://doi.acm.org/10.1145/2591513.2591599

[31] Analog Devices Inc., "ADV7511: 225 mhz, high performance hdmi transmitter with arc," 2017, http://www.analog.com/en/products/audio-video/analoghdmidvi-interfaces/hdmidvi-transmitters/adv7511.html#product-overview.

[32] M. Tanter and M. Fink, "Ultrafast imaging in biomedical ultrasound," *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, vol. 61, no. 1, pp. 102–119, January 2014.

[33] V. Chan and A. Perlas, "Basics of ultrasound imaging," in *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, S. N. Narouze, Ed. New York: Springer-Verlag, 2011, ch. 2, pp. 13–19.

[34] Xilinx Inc., "Best practices for ultraScale power measurement using the KCU105," https://www.xilinx.com/video/fpga/best-practices-ultrascale-power-kcu105.html.

[35] A. Ibrahim, F. Angiolini, M. Arditi, J.-P. Thiran, and G. De Micheli, "Apodization scheme for hardware-efficient beamformer," in *Proceedings of the 12th Conference on PhD Research in Microelectronics and Electronics (PRIME)*, 2016.