Assessment of Image Quality vs. Computation Cost for Different Parameterizations of Ultrasound Imaging Pipelines

Aya Ibrahim[1], Alena Simalatsar[2], Stefanos Skalistis[2], Federico Angiolini[13], Marcel Arditi[3], Jean-Philippe Thiran[3], Giovanni De Micheli[1]

[1] *Integrated Systems Laboratory (LSI), EPFL, Switzerland*
[2] *Rigorous System Design Laboratory (RiSD), EPFL, Switzerland*
[3] *Signal Processing Laboratory 5 (LTS5), EPFL, Switzerland*

ABSTRACT

Ultrasound imaging is a technique widely used in medicine to visualize organs and other body structures, capturing their position, size, morphology and any pathological lesions. Its use is unfortunately limited to specialized centers with trained personnel, and it would be beneficial to expand its applicability to environments like on-the-field emergency response and family physician cabinets. This requires the development of new ultrasound platforms that must be faster, lower-power, easier to use, safe and reliable. One of the major challenges to be met is to dynamically manage a myriad of different imaging options and configuration parameters, which impact image quality and computation cost at the same time. Focusing on this challenge, in this paper we first give an overview of ultrasound imaging techniques and of their possible configuration and parameterization options. We then discuss the impact of these options on computation cost and image quality, showing outcomes from a prototype Matlab ultrasound imaging pipeline.[1]

General Terms

Ultrasound imaging, biomedical imaging, computation cost, image quality, parameterization.

1. INTRODUCTION

Ultrasonic imaging has been playing an important role in medical treatment and diagnostics [15] in a large number of different fields, for example in obstetrics, gastroenterology, cardiology, emergency medicine, etc.. Among the key advantages of ultrasonic imaging are its non-invasiveness, its complete safety compared e.g. X-rays, and its much lower infrastructural complexity compared to other techniques like Magnetic Resonance Imaging (MRI).

Ultrasound imaging is also eminently flexible. The "B-mode" imaging technique is commonly used to visualize organs and inner structures in realtime, capturing their position, size, morphology, and any pathological lesions. Techniques like Doppler imaging [8] on the other hand provide a velocity map, that can be used, for example, to analyze vascularization and helping to identify clots, valve refluxes, or any blood vessel occlusions. Elastography [14] is yet another ultrasound imaging technique; it measures the elastic response of soft tissue to mechanical strain, with the goal of identifying the presence of anomalous tissue - for example, a cyst or tumor - by way of its different elasticity.

Ultrasound scans are performed with a probe, typically hand-held by a sonographer, that is placed directly on a patient's skin. Gel is applied to improve acoustic coupling. The probe emits short bursts of acoustic waves, with a frequency range of 1 MHz to tens of MHz, and a duration of a few microseconds. After each burst, the probe switches to receive mode and performs the reverse operation, capturing echoes reflected by the patient's tissues. Crucially, the excitation waves, which propagate at a speed of about 1540 m/s, are reflected to a different extent by the patient's tissues at acoustic impedance interfaces, which often correspond to tissue density interfaces. Therefore, the delay, intensity and spatial distribution of the echoes is a function of the patient's body structure at different depths. The echoes received by the probe are converted into electrical signals, amplified, sampled and then processed digitally to reconstruct images. The insonification cycle can be sustained at a repetition rate of up to 10 kHz.

The overall imaging process is an exercise in tradeoffs. For example, higher-frequency acoustic waves improve resolution thanks to shorter wavelengths, but incur a loss of penetration, making them suitable only for shallow structures. Focusing the excitation wave into a beam increases the insonification energy along that beam, improving image quality locally, but at the cost of resolution elsewhere in the frame. More sophisticated image processing, e.g. algorithms that compound multiple insonifications in one image to average out noise or maximize resolution, improve the quality of a single frame, but can affect adversely the frame rate. Additionally, each component of the imaging system, including the probe itself, can be configured in many possible ways, resulting in a different balance of performance, image quality, and manufacturing or computation cost.

Despite the relative maturity of the ultrasound imaging market, in recent years several different major trends and innovations have been emerging. For example, traditional "2D" ultrasound techniques, which produce a 2D frame at

a time, have been expanded to "3D" systems, which reconstruct a whole 3D volume at a time. Key advantages include faster acquisition of patient data when large organs must be imaged, easier monitoring of volumetric structures (e.g. heart chambers), and sometimes radically better images, especially in obstetrics. Unfortunately, 3D imaging requires orders of magnitude more calculations and memory than 2D, posing new challenges. Simultaneously, there is a constant push for cheaper, more portable systems, ideally battery-operated, wireless, and even internet-connected. A highly relevant application is imaging where no power supply is available (e.g. on the field or in developing countries). Another scenario is operation by any generic medical personnel untrained in sonography, with a remotely-located experienced sonographer interpreting the images.

In summary, there is a need for the development of scalable, low-power, high-performance ultrasonic platforms, running on sophisticated hardware/software platforms that must meet safety and reliability specifications while accounting for a myriad of different imaging modes and parameters. For example, the software must guarantee that frames are delivered on time at a constant rate. While meeting this constraint, the available time should be used to optimally tune the imaging configuration, based on sonographer directives, based on plugged vs. battery-powered status, or possibly even automatically in a content-dependent way. Providing appropriate software guarantees on functionality and image quality, in presence of such complex imaging options, is a paramount challenge to tackle. As a stepping stone towards this ultimate goal, in this paper we present:

- An overview of ultrasound imaging techniques and computation pipelines.
- A discussion of how an ultrasound system is configured and parameterized, and of how these tradeoffs impact computation and image quality.
- A work-in-progress Matlab toolchain that has been developed to study such tradeoffs.
- A set of case studies to illustrate the effect of imaging options, and to demonstrate our Matlab toolchain in operation.

2. PREVIOUS WORK

Image quality, in the context of medical ultrasound systems, is highly subjective and its characteristics vary among different US image processing applications. If any guarantee is to be provided on the Quality of Results (QoR), it is important to analyze the factors that affect image quality.

In [17], the authors studied the impact, in terms of quality, of Tissue Harmonic Imaging (THI) in contrast with conventional ultrasound methods. The study was focused on imaging organs of the abdominal area, mostly the pancreas, and found that THI yields better image quality, according to experts. A similar study on the liver was conducted on patients with suspected lesions, yielding similar results [4].

A more recent study [23], for solid and cystic lesions of the liver, outlines the differences between THI, spatially compounded sonography, their combination and conventional US imaging. While conventional ultrasound imaging exhibited poor quality compared to the other techniques, none of them outperformed all the others in all scenarios. Spatially compounded sonography had better quality in depicting solid lesions and fatty liver, THI in depicting cystic lesions, and their combination had fewer unwanted artifacts and better overall quality.

In [13], on the other hand, the capability of ultrasound systems to capture dynamic changes are investigated, specifically estimating the amount of blood-flow in arteries around the neck. Again, the authors show that none of the imaging techniques (B-flow, Color Doppler and Power Doppler imaging) consistently outperforms the others, and that each one is suited for different diagnostic purposes.

These are only a few of the many studies evidencing that image quality is highly dependent on i) the insonification method, ii) the ultrasound imaging technique, iii) the object under study, and iv) the suspected condition for which the examination is performed. It is thus imperative for general-purpose platforms to be highly adaptable in order to provide high quality outcomes. In [3], the authors present an automatic image quality adjustment inspired by the autofocus capability of conventional digital cameras, which in case of ultrasound imaging requires simultaneous adjustment of several acquisition parameters (the speed of sound estimate, the focus, the frequency and the mode of operation). The adjustment of parameters is derived from image analytics using a machine learning algorithm and shows an improvement in image quality. Such adaptability may be very helpful for unskilled operators and may be considered in our future work. However, the constraints on computational resources mandate smart management of the parameters affecting image quality and computation cost.

In [12] the authors investigate, identify and discuss the sources of artifacts such as drop-out and shadowing in 3D ultrasound imaging. The dynamic adaptability of image quality, though, requires some objective metrics. Two objective image quality metrics for ultrasound imaging such as Point Spread Function (PSF) and simple Contrast-to-Noise Ratio (CNR) metrics are presented in [19]. These are well-known metrics that can be used to evaluate the image quality, but only if a golden reference for the image is available, i.e. if the scan involves an artificial phantom. In clinical studies the subject's body is not known a priori, and therefore the dynamic adjustment of the image quality based on these metrics is hard.

In this work, we evaluate parameters affecting image quality versus computation cost, which is important when developing a scalable low-power, high-performance and trusted ultrasound platform. In such systems, goals of utmost importance are quality control, feasibility and optimality. Feasibility, in terms of timing, means that no processing task must miss its deadline. While operating under this constraint, the system must make optimal use of its resources and time budget and at the same time provide the best possible image quality. Existing work [2] formulates and addresses the problem of QoS control of real-time multimedia systems in a feasible and optimal manner.

3. ULTRASOUND IMAGING

In order to image a body part with ultrasound techniques, the first step is the transmission of high-frequency (often called RF for Radio Frequency) acoustic waveforms through the region of interest by a process called *insonification*. This is achieved with an ultrasound transducer comprising a set of vibrating elements, often based on the piezoelectric principle. These elements are most commonly physically laid in a straight row (*linear array*), or in a convex arrangement (*convex array*) (Figure 1 left and right). Recently, volumetric or "3D" imaging has also become available. 3D imaging is achieved by means of a probe with a matrix of elements - rather than an array -, or by mechanically sweeping a probe for 2D imaging and by collating the acquired frames.

A linear array has a large probe face. If all elements are

excited in unison, the resulting wave superposition approximates a plane wave. Applying a delay profile to the element excitations allows for more complex wave profiles, most notably a focused beam. In any case, the elements of linear arrays usually feature narrow directivity and the wave can only propagate right in front of the probe. *Phased arrays* are similar to linear arrays, but electronically steer the beam sideways, thus sweeping a circular sector and removing the directivity restriction. Other layouts are also possible for specialized applications, like annular arrays.

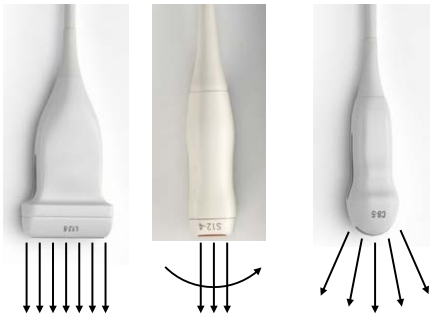


Figure 1: Examples of ultrasound probes and of their acoustic emission. Left: linear array; center: phased array; right: convex array. Probe pictures from the Philips catalog [9].

The acoustic waves are scattered by body structures due to differences in acoustic impedance. Therefore, the body can be seen as a collection of *scatterers* with different reflective properties. The imaging problem is therefore equivalently formulated as the problem of reconstructing the distribution and amplitude of scatterers in space based on the backscattered echoes.

To do so, the returned modulated RF echoes are processed with a set of algorithms as shown in Figure 2. First of all, an amplification followed by Analog-to-Digital Conversion (ADC) is needed to digitize the analog echo signals. It is important to note at this point that ultrasound image reconstruction is bandwidth-intensive. For example, a very simple 64-element linear array for 2D imaging, with 15 MHz center frequency and 10 MHz bandwidth, sampled at 50 MHz to comply with Nyquist-Shannon's theorem [16] and at 12 bits to preserve sufficient dynamic range, generates a stream of data of 4.8 GB/s, that must be processed in realtime. A matrix probe for 3D imaging possesses roughly 100 times more elements, and thus in principle requires 100 times more raw bandwidth; usually, however, since cables with thousands of wires are impractical, either multiplexing or analog preprocessing are employed, bringing the required bandwidth to similar levels again.

Since acoustic waves incur propagation attenuation, the received echo amplitude should be compensated according to depth, which can be realized by increasing the receiver gain over time through a step called Time-Gain Compensation (TGC). It is optional to apply decimation to the signal to reduce the time complexity of the reconstruction algorithms, while following the guidelines of the Nyquist-Shannon sampling theorem. Afterwards, the signal should be high-pass filtered to remove the low-frequency bias that can be due to saturation of amplifiers under some operating conditions.

The most crucial and computationally expensive stage of ultrasonic imaging is *beamforming*. Beamforming constructs an image by applying delays to the returned echoes and summing them together to map the location of their origins. The applied delays on the returned echoes can be seen as a form of focusing applied during receive. We can express the problem of identifying the echo samples that should be summed to focus as follows:

$$s(S) = \sum_{D=1}^{N} w(S)e(D, t_p(O, S, D)), \forall S \in V \qquad (1)$$

S is a point in the volume of interest V. $s(S)$ is the reflectivity of scatterers at location S, which will ultimately be used to calculate the brightness of the corresponding image pixel. N represents the number of receiving elements accessible in the probe, and e is the amplitude of the echo received by element $D \in 1, \cdots N$ at the time sample t_p, which is chosen to correspond to the propagation delay of sound from an origin O to S and back to D. w is a weighting representing the apodization and attenuation to compensate for the inevitable side lobes of the emitted waves [21]. The calculation must be repeated $\forall S \in V$.

The high-frequency beamformed signals should be demodulated to get the baseband information. Two common demodulation techniques used for this purpose are Hilbert transform and IQ demodulation. A logarithmic compression step is then necessary to compress the extremely large dynamic range - up to 100 dB - of ultrasound echoes into a more usable range of gray levels.

The final step of the imaging pipeline is *scan conversion*. Scan conversion is the process of identifying screen image pixels starting from beamformed echo signal amplitudes. Since the latter are not necessarily represented at the same size and aspect ratio of the screen image, upsampling or downsampling may be involved. Furthermore, in phased array or convex array imaging, scan conversion changes the coordinate system from polar to cartesian (see Figure 3).

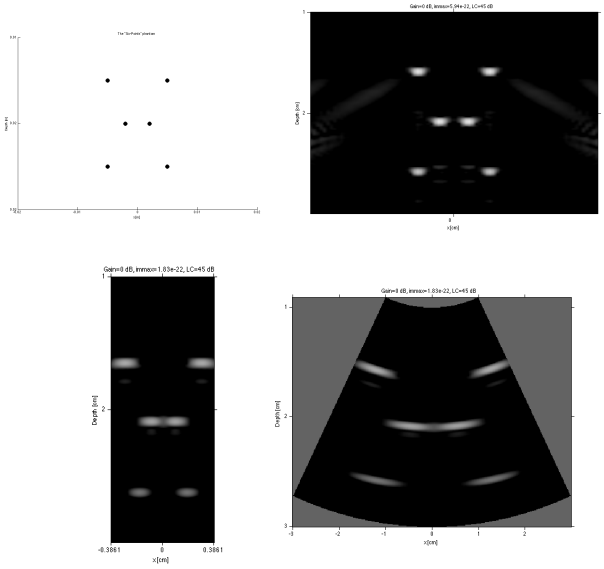
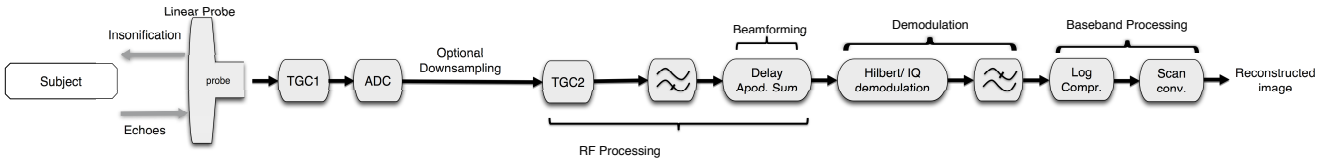
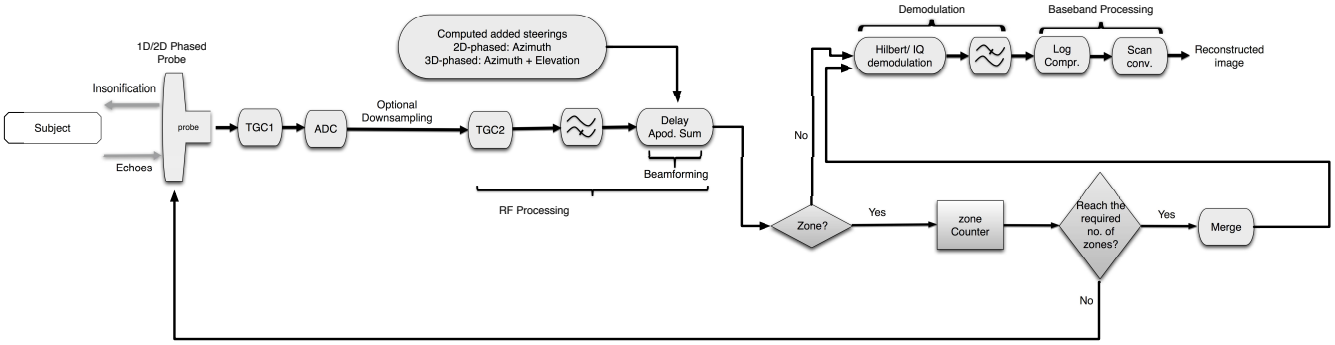


Figure 3: For a synthetic phantom comprising six punctiform scatterers (top left), an image can be reconstructed by simulating a linear array and imaging a rectangular region (top right), or a phased array and imaging a circular sector (bottom). Bottom left: the beamformed sector, still in polar coordinates; bottom right: the same sector after scan conversion.



(a) 2D linear ultrasound imaging pipeline using *1D linear* array transducer.



(b) 2D/3D phased ultrasound imaging pipeline using *1D/2D phased* array transducer.

Figure 2: Ultrasound imaging pipelines.

4. PARAMETERIZATION OF ULTRASOUND IMAGERS

As there is no single way of reconstructing ultrasound images, it is essential to recognize the main factors that affect image quality versus the computation cost according to the specific setups of each clinical exam, and then to optimize such quality within the limits imposed by the available hardware platform.

4.1 Clinical Factors

The reconstruction of an ultrasound image is in large part dependent on the design of the probe used to acquire the raw echo data. In turn, this design is optimized for a given clinical purpose, as can be seen in the wide variety of probes available on manufacturer catalogs [9, 18]. Transducer and cable manufacturing costs also come into the picture, resulting in many possible alternative probes being available in a clinical setting.

For example, linear probes (see Figure 1, left) can usually provide excellent image resolution as these transducers are large enough to house many piezoelectric elements, and these elements being parallel, they can easily emit a plane wave with good pressure uniformity, resulting in homogeneous image quality across the frame. However a linear probe is typically large, and thus unsuitable for imaging in narrow spaces. Another issue due to the linear geometry and size is to always maintain a good contact - and thus, acoustic coupling - with the patient's skin across the whole surface.

Better contact can be achieved with small-footprint convex arrays or phased arrays (see Figure 1, right and center). The probe thus can be more easily pressed on the skin. The ability to image large volumes is recovered either thanks to the convex design, or in phased arrays, with a piezoelectric tuning optimized for broad directivity of each element's acoustic emission. These arrays are a much better option for imaging in tight spaces, e.g. intercostally (for cardiac imaging), or inside body cavities, e.g. intra-abdominally. Unfortunately, since they are limited to smaller apertures, their lateral resolution is affected. The optimal insonification generated by both convex and phased arrays covers a sector. The wider imaging area at the far end of the sector means that acoustic pressure decreases very noticeably with depth, and resolution is correspondingly degraded at the deep end.

Another crucial factor driven by clinical considerations is the probe's acoustic frequency. Higher frequencies allow for better feature resolution, but are attenuated more by the patient's tissues, with a loss of penetration. The piezoelectric elements of a transducer must be fabricated for optimum resonance at a given frequency, and although wideband probe design has been an active research area for a long time [1], most probes can only operate in a limited range around the optimum frequency. Typical pulse frequencies for medical ultrasound can range between 2 and 30 MHz, with the lower end of the spectrum allowing for imaging up to 20-30 cm deep (e.g. for abdominal scans) and the higher end being optimized for high-resolution superficial imaging, e.g. for ophthalmology.

Since the clinical application is not under the control of the imaging system, all the parameters related to the chosen probe should be considered as constraints by the imager designer. The key ones from a quality/computation cost standpoint are:

- *Number of the elements*, and of the elements that can be simultaneously active (function of the cabling). This parameter is crucial for image quality but impacts linearly the computation cost per pixel.
- *Physical arrangement* of the elements, e.g. linear or convex. This parameter impacts the beamforming stage of the imaging pipeline.
- *Directivity* of the elements (function of their size) and thus suitability to beam steering. This parameter also impacts the availability of processing options.
- *Acoustic frequency* of the probe (function of the chosen materials and sizes). This parameter linearly impacts the computation cost, although high frequencies result in shallower images, which may require the computation of fewer pixels.

4.2 Imaging Mode

To reconstruct an ultrasound image, multiple techniques and optimizations are available. Some of them are dictated by the choice of the probe, but others remain under the full control of the designer of the imager. These include:

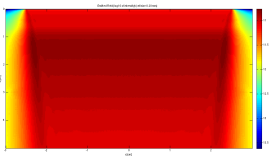
- *Transmit focusing.* Any probe can be driven so that its elements vibrate in unison, or with a certain delay profile. The overall acoustic wave changes in wavefront and direction of propagation, as seen in Figure 4. It is possible to distribute the acoustic pressure evenly in the volume of interest, or to concentrate it in specific areas, yielding better local resolution. It is even possible to choose a diverging beam, where a virtual focal point is behind the transducer, to insonify a wider region. Transmit focusing has a quality impact on the images, but no particular overhead on computation times. It complements the receive focusing that is done in beamforming.

- *Insonification pattern.* The simplest approach is to reconstruct an image frame from each insonification. Ultrahigh frame rate imaging [10] relies heavily on this approach. Noting that acoustic waves in the body travel at approximately 1540 m/s, the two-way propagation delay, for a 20 cm imaging depth, is of about 260 μs, meaning that with quick-enough signal processing, upwards of almost 4000 frames per second could in theory be produced. However, this level of speed requires challenging electronic design, and is downright impractical, with current technology, for 2D even more so for 3D imaging.

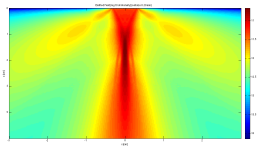
  An alternative is *zone imaging*, whereby a single frame is reconstructed from multiple insonifications. The volume of interest is divided in zones - for example, in circular subsectors -, each of which is sequentially beamformed based on a new insonification (Figure 5). Each new insonification may correspondingly shift the focus location, optimizing resolution in that zone. The zones are then stitched into a frame. This approach reduces the maximum possible frame rate by a factor equal to the zone count, but optimizes intra-frame resolution, and is better suited to systems where computation is slow and the cabling constrains the number of simultaneously active elements in receive - such as in 3D imaging. A downside of zone imaging is that patients can never be perfectly still, resulting in potentially significant alignment glitches across zones.

- *Frame compounding.* Any acquired ultrasound image frame is bound to have sub-optimal resolution and feature a certain degree of adverse phenomena including speckle, due to interference among acoustic waves and multiple reflections, and noise. This can be partially remediated with compounding techniques, whereby a single frame is constructed from multiple ones by techniques like blending or averaging. This comes at a corresponding penalty to frame rate, and is a good match to ultrahigh frame rate imaging, where the availability of such high throughput can be exploited to improve frame quality.

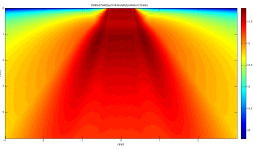
- *Harmonic imaging.* The acoustic propagation in the body is a non-linear process, giving rise to reflections that are harmonics of the insonification frequency; usually only the first harmonic has a significant amplitude. Provided that the probe has sufficient bandwidth to detect echoes at frequencies at least twice as high as the transmit frequency, the imager can then process the received echoes demodulating from the first harmonic of the nominal frequency, yielding better resolution.



(a) Plane wave, using 192 elements probe



(b) Converging wave, using 16 elements probe



(c) Diverging wave, using 16 elements probe

Figure 4: Different transmit focusing options. The transducer is at the top of the figures.

4.3 Imaging Parameters

Once a suitable imaging mode has been identified, it is still possible to tune a certain number of signal processing parameters that dramatically affect the image quality. For example, if not all probe elements can be read out simultaneously due to hardware restriction, the specific pattern of which elements are read out can be chosen dynamically. As another example, towards the end of the imaging flow, the image brightness can be tuned to ensure optimum feature visibility; this is usually done on a logarithmic scale since the dynamic range of the echoes exceeds 100 dB. These two configuration options, and many others, are either computationally cheap, or they incur an approximately fixed runtime cost. We thus emphasize in this paper two main parameters that have both a major effect on image quality and runtime:

- *Sampling frequency and decimation.* According to the Nyquist-Shannon theorem, the received echoes should be sampled at a frequency at least twice the transmit frequency, in order not to lose information content. Therefore, the sampling frequency of an ultrasound system is partially fixed by the chosen probe's maximum frequency. Unfortunately, the sampling frequency linearly affects the bandwidth of the input samples to be beamformed, and must not be exceedingly high. A technique called *decimation* can optionally be employed on the echo signal samples, with a similar effect to the use of a lower sampling frequency. By accepting a quality loss, it is even possible to use decimation factors that violate the Nyquist constraint, achieving a further computation speedup.

- *Image resolution.* The amount of pixels beamformed by the imager linearly determines the runtime of the beamforming routine. Thus, it is possible to control the computational cost with a different choice of resolution of the output image. The ideal resolution should be chosen to match the resolving power achievable by the probe/imager combination, but this latter is hard to quantify. Additionally, some clinical applications demand the highest possible resolution while attaching secondary importance to frame rate (e.g. tumor or small-feature detection), others demand just the opposite (e.g. measurement of the blood velocity in vessels). Thus, in practice, the resolution can be chosen depending on the desired quality/frame rate tradeoff.

- *Active element pattern in receive.* To reduce the amount of data to process, it is possible to beamform

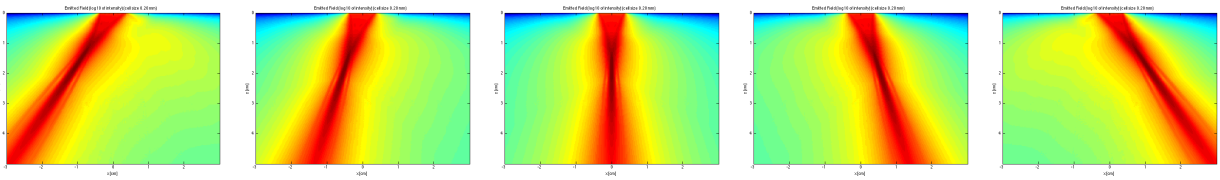


Figure 5: Example of possible focused beams for zone imaging, where five sectors are insonified in sequence.

based on a sub-aperture of the whole probe. In other words, it is possible to switch off some probe elements in receive, and to perform reconstruction based on the echoes received by the remaining active elements only. Naturally, this degrades image quality. A probe with sparser elements features a radiation profile with a smaller main lobe and larger side lobes, called *grating lobes* if they become comparable in size to the central lobe. Grating lobes severely affect the directional selectivity of the probe, leading to artifacts that can be major. It is therefore important to choose a sub-aperture that reduces calculation requirements within an acceptable quality degradation.

4.4 Evaluation of Ultrasound Image Quality

The evaluation of ultrasound image quality is a complex and largely subjective task.

A first key reason is that different clinical exams require different priorities - for example, resolution is critical for tumor detection, while frame rate is more important for the study of dynamically changing tissue, such as cardiac imaging.

Furthermore, even when the reconstruction goals are clear, their measurement and optimization is still challenging. For example, it is generally accepted that "high spatial resolution" (defined as the ability to distinguish nearby scatterers) and "high contrast resolution" (defined as the ability to distinguish scatterers of different reflectivity) are positive image traits, while "speckle" (granular image artifacts due to the coherent nature of acoustic beams) and "noise" (any degradation of the actual signal due to imperfect acoustics and electronics) are to be minimized. Unfortunately, three main issues arise.

- These metrics are either loosely defined or have multiple, slightly different definitions in literature, see for example [20] for contrast resolution.
- It is very challenging to measure most of these parameters with a real system *in vivo*, simply because there is no golden reference image available. Instead, the clinician often relies on his experience to subjectively evaluate if the image is "clean" or not. To work around the issue, *phantoms* are often used. A phantom is a physical object composed of materials with different reflectivity, and a precise internal structure designed to stress the imager's reconstruction capabilities. In the context of a Matlab simulation environment, a phantom is a synthetically-defined geometrical structure that is to be insonified and imaged. Still, there is no commonly accepted way to quantify how invasive speckle is, and electronic noise cannot be modeled in a general way in this domain unless a very precise reference system is first built and characterized.
- Although it is possible, to a degree, to boost the positive image traits and to minimize the negative ones, these two opportunities often compete against each other. For example, noise and speckle can be largely reduced, but only at a price in spatial and contrast resolution, or vice versa. This makes it futile to attempt a precise, invariant characterization of image quality. In practice, a real-world system is likely to have knobs that the sonographer operates based on his current needs, maybe based on a specific part of the image alone that is of higher clinical relevance.

A commonly cited [19] quantitative image quality metric is the Point Spread Function (PSF), defined as the output produced by the ultrasound system when imaging an ideal object comprising a single reflective point in space. It can be shown that the Fourier transform of this function is the transfer function of the overall system, and thus, once known, it is possible to predict the output image generated for any arbitrary set of scatterers by way of a convolution of such input with the PSF. A PSF should ideally be infinitely compact, although in practice its compactness is limited by diffraction laws and by the finite duration of the acoustic pulse used for the insonification. This bounds the achievable spatial and contrast resolution.

Even though the PSF can be useful, it is still very complex to provide a quality assessment based on it, since the PSF typically varies substantially depending on the location in the image, and since it has distinct axial and lateral components. The PSF is also influenced by numerous parameters of the ultrasound probe and of the imaging pipeline.

For all the reasons above, in this paper we will abstain from a detailed quantitative quality assessment, and will instead focus on the more obvious and qualitative dependencies that can be observed when changing the main parameters of the imager. As will be seen in the next Section, it is still possible to subjectively discern differences, sometimes major differences, as a function of the imaging parameters.

5. A MATLAB ULTRASOUND IMAGING KIT

Endeavouring to design a powerful and flexible ultrasound imaging system, development and testing in an environment such as Matlab is essential. This is needed to assess the most important algorithmic alternatives and their impact on the image quality and computation cost. Only later the design can be mapped onto a hardware/software platform, for example running off one or more FPGAs. Even then, a simulation environment is still useful for debugging, for the generation of specialized inputs, for the evaluation of the outputs, and for the exploration of new features. Therefore, we have developed a feature-rich Matlab environment that models complete imaging pipelines, like those of Figure 2.

Our Matlab environment has two different options for the generation of input data. The first option is the acquisition of real ultrasound data in realtime from the ULA-OP device [22], [11]. ULA-OP is a powerful and complete platform for ultrasound research purposes. A limitation of ULA-OP is that it only supports probes for 2D imaging.

A more flexible alternative is the generation of synthetic phantoms directly within the Matlab environment. A synthetic phantom is a defined collection of scatterers in space, with configurable reflectivity. The advantage of synthetic phantoms is that they can be arbitrarily configured and that they provide a golden reference for the imaging pipeline, allowing for objective comparisons between the input structure and the final image, which is not possible when using real ultrasound data. We have created a small database of 3D phantoms for development and testing purposes. The free Field II software [6], [7] can then be used to simulate the insonification of this phantom. Field II runs on Matlab; it simulates both emitted fields and pulse-echo fields, and it allows for the definition of various types of transducers, including models of probes for 3D imaging that otherwise may be too expensive to acquire or unavailable for research purposes. The output of a Field II simulation is a set of pressure signals sampled in time. It should be noted that a Field II simulation, while highly accurate and representative, does not model some of the phenomena occurring during a real insonification; for example, it does not generate harmonics of the trasmitted pulse, it does not model multiple reflections, it does not directly model variance in the speed of sound in different tissues, and it does not factor in any electronic or acoustic noise. For any study where these effects are of particular concern, ULA-OP can be used.

We have implemented the imaging pipelines shown in Figure 2, covering both 2D and 3D imaging. Some of the major implemented imaging modes, as will be shown in Section 5.1, include:

- Using linear or phased array probes, with transmit focus that can be a plane wave or a focused beam, refer to Figure 3. Diverging beams and other focus patterns are also allowed.
- Using a single insonification per frame versus zone imaging, refer to Figure 8.
- Using beamforming with the traditional delay-and-sum approach, see Equation 1, or with a built-in Field II function for reference. For 2D and 3D phased array imaging, we support a delay table steering method described in [5].
- Envelope extraction by either IQ demodulation or Hilbert transform.

5.1 Case Studies on Pipeline Parameterizations

In this section, we show several experiments using the Matlab toolchain we have developed. The goal is to present an overview of how imaging parameters affect the subjective quality of the results, while at the same time estimating the computation cost via the runtime of the Matlab simulation. We do not aim to identify configurations that are "best", because each of them is likely to be useful in some scenarios, but rather to assess the quality/computation trade-off, in view of a final implementation able to switch modes on-the-fly based on the available energy supply and on the currently displayed image.

The computation cost is evaluated by measuring the runtime of the Matlab code that is in charge of image reconstruction (i.e., excluding the runtime of the Field II simulation, which is irrelevant for our purposes). We used Matlab R2012b on a 2.4 GHz Intel Core i5 with 4 GB of DDR3 RAM at 1600 MHz, averaging three runs of each experiment. Obviously, these execution times give a very rough estimation of the actual computation costs. However, in our experience, we have generally noticed a good correlation between variations of Matlab runtime and the projected amount of software calculations or hardware resources that would be needed in the corresponding imaging scenarios. Thus, we consider Matlab runtime a good proxy to estimate how various parameters influence performance, at least in relative terms.

5.1.1 Decimation of Input Samples and Envelope Extraction

This experiment shows the image quality degradation as we vary the bandwidth of the input signal by way of decimation. Additionally, we toggle the method of envelope extraction, using either IQ demodulation or the Hilbert transform.

We insonify with Field II a synthetic phantom consisting of a hollow sphere, of outer diameter 2 cm, crossed by a wire. For this experiment, we model the ultrasound probe as a 96-element linear array with central frequency f_u = 3.5 MHz and bandwidth also equal to 3.5 MHz, yielding a spectrum with a maximum frequency of 5.25 MHz. The echoes are sampled at f_s = 200 MHz. According to the Nyquist sampling theorem, the minimum frequency at which we have to sample the echoes is f_{min} = 10.5 MHz. The maximum decimation factor can be computed as follows: $df = floor(f_s/f_{min})$, yielding df = 19 for our case. This means discarding 18 echo samples every 19.

We present ten reconstructed images in Figure 6. The upper row corresponds to envelope reconstruction using the Hilbert transform, while the lower row uses IQ demodulation. In each row, the decimation factor df is swept from 1 (no decimation applied) to 25 (in violation of the Nyquist constraint). As expected, the image quality is progressively degraded with the increase of the decimation factor, with the appearance of speckle. However, the phantom is still imaged with good quality up to $df = 7$, with reasonable quality up to $df = 16$, and remains easily recognizable even upon extreme decimation. This surprising result indicates that in severely battery-limited emergency situations, aggressive decimation could be used as a measure of last resort.

The runtimes corresponding to these same experiments are presented in Table 1. Since the decimation factor linearly impacts the beamforming time, which dominates the overall reconstruction cost, the decimation factor is almost directly proportional to the latter.

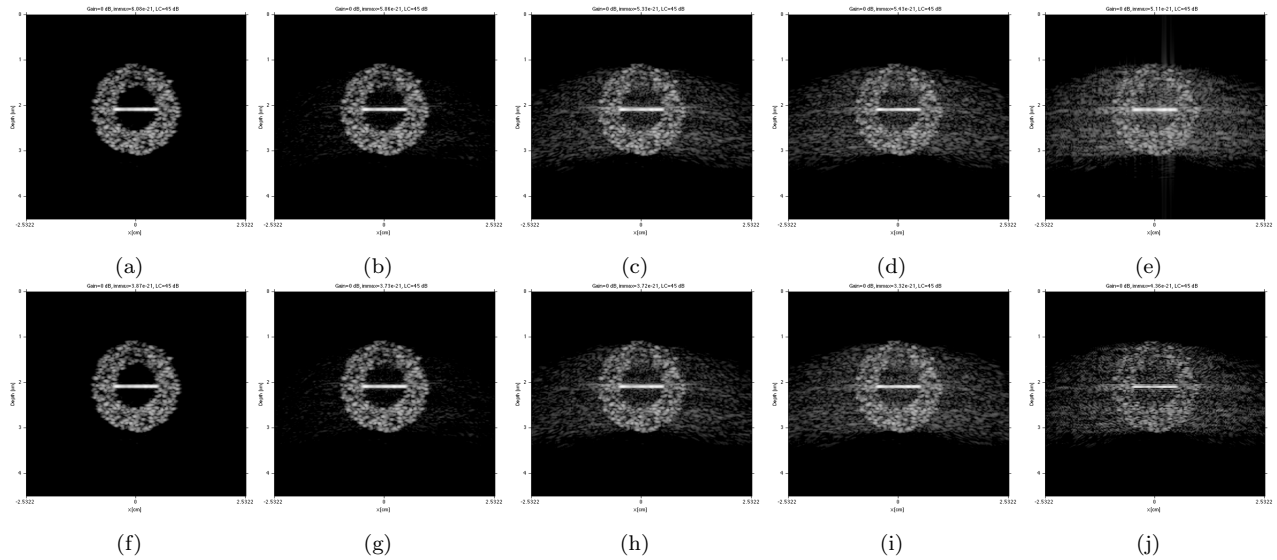
The two alternatives for envelope extraction perform very similarly and are essentially indistinguishable until $df = 19$. At $df = 25$, IQ demodulation seems to preserve more resolution. In this case, the results suggest that the lowest-overhead technique - i.e. IQ demodulation - is the better choice.

Table 1: Decimation factor vs. execution time

Execution Time	df=1	df=7	df=16	df=19	df=25
Hilbert	9.9 s	2.7 s	1.7 s	1.7 s	1.5 s
IQ	9.6 s	2.6 s	1.8 s	1.6 s	1.5 s

5.1.2 Active Element Pattern in Receive

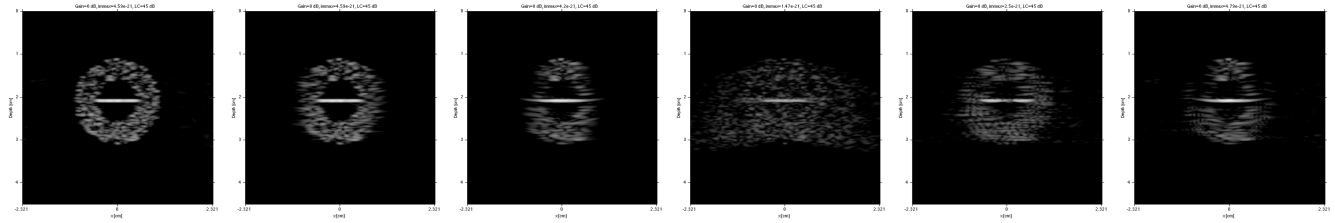
In this experiment, for the same 96-element linear array presented above, we investigate the effect of choosing different activation patterns of the elements, i.e. keeping some of the 96 elements inactive. Note that it does not make sense to do so in transmit, where a plane wave is emitted at no particular computation cost; in receive, however, the beamforming



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Figure 6: Image reconstruction of a hollow spherical phantom with a 96-element linear probe, having f_u=3.5 MHz, f_s=200 MHz. Upper row: envelope reconstruction with the Hilbert transform; lower row: with IQ demodulation. Left to right: sampled echoes decimated with df=1 (no decimation), 7, 16, 19 (Nyquist threshold), 25.



(a) 96 out of 96 (b) 32 out of 96 (c) 16 out of 96 (d) 1 every 3 (e) 6 every 16 (f) 12 every 32

Figure 7: Image quality depends on the number and pattern of probe elements that are *active* in receive. Tests on a 96-element linear array, with full aperture used in transmit.

cost can be reduced proportionally to the number of inactive elements. We report the obtained images in Figure 7, and the corresponding runtime in the top rows of Table 2.

Figure 7a shows the best possible image quality, when all 96 elements are active. Figures 7b and 7c show reconstruction results with a smaller receive aperture, by using only the central 32 and 16 elements respectively. A reduced aperture always results in a loss of lateral resolution and, in the extreme case of using only 16 elements, also reduces the ability to image the edges of the phantom, which are only partially insonified by the narrower emitted wave. A "naïve" activation pattern where only one element every three is active (Figure 7d) does insonify the whole volume, but result in unacceptable speckle due to the emergence of *grating lobes*, as discussed in Section 4.3. Grating lobes increase in intensity proportionally to the spacing of the active elements. Another possible arrangement is to use multiple clusters of active elements, namely 6 every 16 elements (Figure 7e) or 12 every 32 (Figure 7f), in both cases with a total active element count of 36. These arrangements incur less speckle than using 1 active element every 3, while saving a comparable amount of calculations; however, their coverage of the volume and lateral resolution are more compromised. Note, as seen in Table 2, that computation cost only depends on the active element count, not on the pattern.

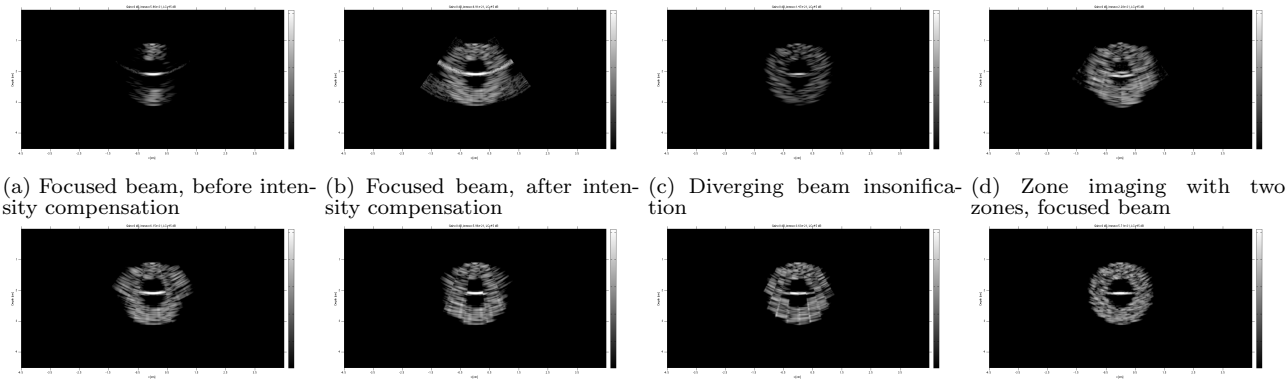
Overall, depending on the clinical requirement on the width of the region to be imaged, different trade-offs are available; if a broad region is being analyzed, a full aperture is the most sensible choice, but if only a thinner sector is of actual concern, a smaller aperture may represent a more efficient compromise.

Table 2: Runtime of the imaging flow, for different array types and element counts.

Array	TX focus	RX active	Avg runtime
Linear	Plane	96/96	9.63 s
Linear	Plane	32/96	6.02 s
Linear	Plane	16/96	4.8 s
Linear	Plane	1/3 ×32	6.02 s
Linear	Plane	6/16 ×6	6.36 s
Linear	Plane	12/32 ×3	6.36 s
Phased	Focused	32	8.12 s
Phased	Focused + brightness comp.	32	8.17 s
Phased	Diverging	32	9.9 s
Phased	Zone imaging	32	8.17 s

5.1.3 Phased arrays and zone imaging

We now present experiments with a simulated 32-element phased array, still with f_u=3.5 MHz. Figure 8a shows the reconstructed phantom when using a focused beam in transmit. A focused beam yields maximum resolution in the fo-



(a) Focused beam, before intensity compensation (b) Focused beam, after intensity compensation (c) Diverging beam insonification (d) Zone imaging with two zones, focused beam

(e) Zone imaging with three zones, focused beam (f) Zone imaging with five zones, focused beam (g) Zone imaging with ten zones, focused beam (h) Zone imaging with fifty zones, focused beam

Figure 8: Phased array imaging with emission profiles including focused and diverging beams. (a)-(c): single insonification per frame; (d)-(h): zone imaging with 2-50 zones, each being a circular sector.

cused area, since the acoustic pressure is highest, but since less energy is used to insonify the rest of the volume, the unfocused regions exhibit lower resolution and brightness. The latter effect can be compensated very cheaply by applying a brightness adjustment profile (Figure 8b). Alternatively, if a diverging beam is used in transmit (Figure 8c), the volume can be insonified more evenly, resulting in a more consistent frame quality. The choice depends mainly on the clinical interest of identifying small features in a precise area, or rather of studying a broad volume.

We also show images from a zone imaging technique, whereby a single frame is reconstructed by stitching multiple sub-frames, each computed from a separate insonification with optimized focus. In Figures 8d-8h the resulting frame is a composite of 2-50 zones, each being a thin circular sector. For illustrative purposes, in this set of figures we do not apply any blending across the zones, which should normally be done. Therefore, it is easily possible to detect abrupt transitions and artifacts at the edges of the zones, until the number of zones becomes very high (Figure 8h). Increasing the number of zones also helps resolution since a larger percentage of the frame is in focus.

From the computation standpoint (bottom rows of Table 2), it can be seen that the runtime does not vary much; indeed, the total number of beamformed pixels per frame remains the same regardless of transmission focus or zone imaging choices. Brightness compensation and zone stitching are both very cheap operations and they do not skew the results substantially. On the other hand, it is important to remember that in the case of zone imaging, the number of insonifications to reconstruct one frame is equal to the number of zones. Therefore, even if the beamforming workload per frame remains constant, the maximum achievable frame rate may be severely impacted.

5.1.4 3D Imaging and Image Resolution

Imaging runtime is of course highly dependent on the resolution of the beamformed images. The computation cost can be drastically reduced if the image resolution is lowered. We confirm this prediction by simulating a 16x16 element matrix phased array for 3D imaging, and varying the axial resolution of the computed images. The resulting images are shown in Figure 9, and the runtime in Table 3. Of course, lower resolutions degrade the image quality, although not as much as could be expected, even at extremely low values. It is impossible to devise an ideal resolution value in a general way; the guiding principle should be to match the physically achievable spatial resolution, which depends on the probe's element count, element arrangement, central frequency, etc. Note that in 3D imaging, runtimes do not depend linearly on the axial resolution of the beamformed image, as could be expected. This is because we implement a delay table steering mechanism [5] which saves memory but incurs a fixed computation overhead which is not resolution-dependent.

Table 3: 3D beamforming runtime as a function of the desired axial resolution. 11688 is the maximum possible resolution, by reconstructing a pixel for every time sample of the input echoes, considering the f_s and imaging depth of this experiment.

Axial Resolution	11688 pixels	500 pixels	100 pixels	50 pixels
Runtime	61.7mins	15.2mins	13.6mins	13.3mins

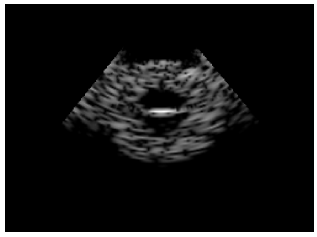
3D beamforming is of course also heavily dependent on the number of elements of the probe. In Figure 10 we show the reconstruction quality of 3D images when using 16x16 and 32x32 matrix phased arrays. The beamforming runtimes in Matlab are 15.2 and 35.6 minutes respectively; ignoring the delay table steering overhead, the runtime of the core beamforming routine is found to be quadratically proportional to the number of active elements in the probe, as expected in theory.

6. CONCLUSIONS

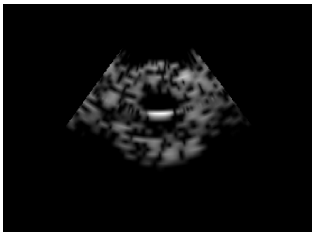
In this paper, we have presented one of the key challenges in the development of a full-featured next-generation ultrasound imager, namely the issue of supporting a large set of imaging modes and parameters, each with a different set of trade-offs in image quality vs. computation time. The challenge is made more complex by the need to evaluate the resulting images in a way that is largely subjective.

We have shown a work-in-progress Matlab toolchain developed at EPFL, that already today supports a large number of the possible permutations of imaging modes and parameters. We have shown numerous examples of images obtained in each of those modes, and commented on the merits of each option. The Matlab toolchain allows us to image a synthetically-defined phantom, which has the advantage of representing a golden reference, but an option also exists to process real-world ultrasound signals.

(a) 500 points reconstructed along the depth axis



(b) 100 points reconstructed along the depth axis



(c) 50 points reconstructed along the depth axis

Figure 9: 3D imaging with a 16x16 matrix phased array, and a variable number of beamformed points in the axial direction. The images show the middle transverse plane of the reconstructed 3D volume.

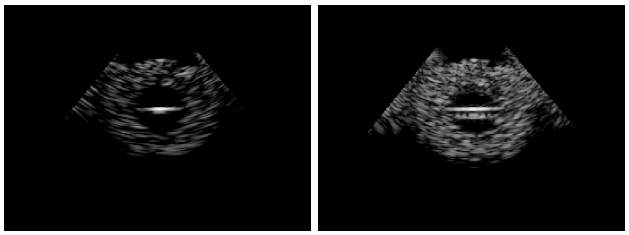


Figure 10: 3D imaging with 16x16 (left) and 32x32 (right) matrix phased arrays. The images show the middle transverse plane of the reconstructed 3D volume.

Work is now underway towards a hardware/software implementation of a 2D/3D ultrasound imager. We are proceeding piecewise towards the mapping of different portions of the Matlab pipeline onto FPGAs and/or software, with the goal of a demonstrator by 2017. The present Matlab codebase is being continually expanded with new imaging options, with the goal of devising alternatives and roughly quantifying computation costs before a hardware implementation is attempted.

Acknowledgments

The authors would like to acknowledge the helpful assistance of Marcel Arditi on the physics of ultrasound imaging. The research work of this paper was funded by the Swiss Confederation through the UltrasoundToGo project of the Nano-Tera.ch initiative.

7. REFERENCES

[1] A. Brown and J. Weight. Generation and reception of wideband ultrasound. *Ultrasonics*, 12(4):161–167, 1974.

[2] J. Combaz, J.-C. Fernandez, T. Lepley, and J. Sifakis. Qos control for optimality and safety. In *Proceedings of the 5th ACM international conference on Embedded software*, pages 90–99. ACM, 2005.

[3] N. Y. El-Zehiry, M. Yan, S. Good, T. Fang, S. K. Zhou, and L. Grady. Learning the manifold of quality ultrasound acquisition. In K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, editors, *MICCAI (1)*, volume 8149 of *Lecture Notes in Computer Science*, pages 122–130. Springer, 2013.

[4] L. Hann, A. Bach, L. Cramer, D. Siegel, H. Yoo, and R. Garcia. Hepatic sonography: comparison of tissue harmonic and standard sonography techniques. *AJR. American journal of roentgenology*, 173(1):201–206, 1999.

[5] A. Ibrahim, P. Hager, A. Bartolini, F. Angiolini, M. Arditi, L. Benini, and G. De Micheli. Tackling the bottleneck of delay tables in 3d ultrasound imaging. In *DATE Conference, 2015. Proceedings of the*, March 2015.

[6] J. Jensen. Field: A program for simulating ultrasound systems. In *the 10th Nordic-Baltic Conference on Biomedical Imaging in Medical and Biological Engineering and Computing*, volume 34, pages 351–353, 1996.

[7] J. Jensen and N. B. Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 39(2):262 – 267, 1992.

[8] Z. Keshavarz-Motamed, J. Garcia, E. Gaillard, R. Capoulade, F. Le Ven, G. Cloutier, L. Kadem, and P. Pibarot. Non-invasive determination of left ventricular workload in patients with aortic stenosis using magnetic resonance imaging and doppler echocardiography. *PLoS One*, 9(1), 2014.

[9] Koninklijke Philips N.V. Ultrasound transducers, 2015. `http://www.healthcare.philips.com/main/products/ultrasound/transducers/open_article_index.wpd`.

[10] J.-y. Lu. Experimental study of high frame rate imaging with limited diffraction beams. *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on*, 45(1):84–97, Jan 1998.

[11] MSD lab, University of Florence. Ula-op system, 2013. `http://www.msdlab.dinfo.unifi.it/CMpro-v-p-19.html`.

[12] T. Nelson, D. Pretorius, A. Hull, M. Riccabona, M. Sklansky, and G. James. Sources and impact of artifacts on clinical three-dimensional ultrasound imaging. *Ultrasound Obstet Gynecol*, 16(4):374–383, Sept 2000.

[13] S. Oktar, C. Yücel, D. Karaosmanoglu, K. Akkan, H. Ozdemir, N. Tokgoz, and T. Tali. Blood-flow volume quantification in internal carotid and vertebral arteries: comparison of 3 different ultrasound techniques with phase-contrast mr imaging. *American journal of neuroradiology*, 27(2):363–369, 2006.

[14] J. Ophir, I. Cespedes, H. Ponnekanti, Y. Yazdi, and X. Li. Elastography: a quantitative method for imaging the elasticity of biological tissues. *Ultrasonic imaging*, 13(2):111–134, 1991.

[15] S. H. C. Ortiz, T. Chiu, and M. D. Fox. Ultrasound image enhancement: A review. *Biomedical Signal Processing and Control*, 7(5):419 – 428, 2012.

[16] C. Shannon. Communications in the presence of noise. In *Proceedings of the IRE,*, volume 37, pages 10–21, 1949.

[17] R. S. Shapiro, J. Wagreich, R. Parsons, A. Stancato-Pasik, H.-C. Yeh, and R. Lao. Tissue harmonic imaging sonography: evaluation of image quality compared with conventional sonography. *AJR. American journal of roentgenology*, 171(5):1203–1206, 1998.

[18] Siemens AG. Ultrasound transducer catalog, 2015. `http://www.healthcare.siemens.com/ultrasound/ultrasound-transducer-catalog`.

[19] C. Simpson. Objective image quality metrics for ultrasound imaging, June 2009. Master thesis.

[20] J. Thijssen, M. Van Wijk, and M. Cuypers. Performance testing of medical echo/doppler equipment. *European journal of ultrasound*, 15(3):151–164, 2002.

[21] K. Thomenius. Evolution of ultrasound beamformers. In *Ultrasonics Symposium, 1996. Proceedings., 1996 IEEE*, volume 2, pages 1615–1622 vol.2, Nov 1996.

[22] P. Tortoli, L. Bassi, E. Boni, A. Dallai, F. Guidi, and S. Ricci. Ula-op: An advanced open platform for ultrasound research. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(10):2207 – 2216, 2009.

[23] C.-L. Yen, C.-M. Jeng, and S.-S. Yang. The benefits of comparing conventional sonography, real-time spatial compound sonography, tissue harmonic sonography, and tissue harmonic compound sonography of hepatic lesions. *Clinical imaging*, 32(1):11–15, 2008.