PROCEEDINGS of the
23rd International Congress on Acoustics
9 to 13 September 2019 in Aachen, Germany

3D imaging method for an air-coupled 40 kHz ultrasound phased-array

Gianni ALLEVATO, Jan HINRICHS, Dominik GROSSKURTH, Matthias RUTSCH, Jan ADLER, Axel JÄGER, Marius PESAVENTO and Mario KUPNIK

Technische Universtität Darmstadt, Germany, allevato@must.tu-darmstadt.de

Abstract

We present a three-dimensional far-field imaging method using an air-coupled 40 kHz 8×8 ultrasound phased-array. A waveguide is attached to the array which reduces the aperture to 35 mm $\times$ 35 mm and the inter-element spacing to $\lambda/2$ for grating lobe free beamforming. Using transmit beamforming, a pulse echo detection is performed sequentially for each direction of the discretized region of interest. All transducers are used for pulse transmission and echo reception. The echo signals are FIR bandpass filtered and sinc-interpolated. The conventional Delay-and-Sum beamformer generates a spatially filtered and summed echo signal. The envelopes of the summed echo signals, extracted by Hilbert filters, are used for generating B-Scans and 3D C-Scans. The signal processing and image generation is performed by a GPU using OpenGL. The imaging properties are characterized in an anechoic chamber. The localization of objects is validated for a range of 0.5 to 6 m in a vertical and horizontal field of view of $\pm50°$. The system is able to detect sound-absorbing objects and objects hidden by obstacles. The implementation has limitations in frame rate and lateral resolution. In this context, future optimization potentials are outlined.

Keywords: Air-coupled, Phased-Array, Imaging

1 INTRODUCTION

In the age of automation, sensors for 3D environmental monitoring are gaining increasing importance. These sensors are typically based on LIDAR, RADAR and camera systems. Ultrasound phased-arrays for 3D imaging are well established in medical imaging, sonar and NDT applications. However, automotive and robotic applications usually require air-coupled ultrasound.

We present a 3D imaging method based on an air-coupled 64 channel ultrasound phased-array. In order to avoid strong sound attenuation by the medium, we use efficient transducers at 40 kHz. As the dimensions of the transducers are too large ($\varnothing$10 mm) to form an array that meets the $\lambda/2$ criterion [1], we use a waveguide [2] based on [3, 4] that reduces the inter-element spacing. This makes grating lobe free beamforming possible [5, 6].

The ultrasound-based method presented is a robust, flexible solution with low hardware requirements, which complements the previously mentioned sensor systems.

2 SYSTEM OVERVIEW

The phased-array system consists of four main components, which are the phased-array itself, the analog electronics, an FPGA board with a Zynq 7010 System-on-Chip and a conventional PC (Figure 1).

The phased-array uses $M = 64$ ultrasound transducers (MA40S4S, Murata, Tokyo, Japan), with a resonant frequency of 40 kHz. The transducers are arranged in an 8×8 uniform rectangular layout and are mounted on a 3D-printed waveguide with 64 sound channels. This reduces the effective inter-element spacing to half the wavelength ($\lambda/2 = 4.3$ mm) and the effective aperture to 35 mm $\times$ 35 mm (Figure 2).

The analog electronics is partitioned into eight identical subgroups. Each subgroup is dedicated to eight ultrasound transducers and contains one transducer pulser IC (HV7355, Microchip, Chandler, AZ, USA), two TR-switches (TX810, Texas Instruments, Dallas, TX, USA), two variable gain amplifiers (AD8335, Analog Devices, Norwood, MA, USA) and one ADC (AD7761, Analog Devices, Norwood, MA, USA). The pulser IC

receives a serial logic signal and generates eight controllable unipolar 40 kHz square-wave signals with an amplitude of 20 V_{pp}. These signals drive the ultrasound transducers. The TR-switches are used to switch between transmitting and receiving electronics. With the VGA, the received signals of the transducer can be amplified variably. Thus, the air attenuation of distant echoes can be compensated. The ADC digitizes the signals in parallel with a sampling rate of 200 kHz per channel.

The Zynq 7010 SoC (Xilinx, San Jose, CA, USA) includes an FPGA and microcontroller in one package. The FPGA is used to control the pulse echo sequence and to read the data of all ADCs in parallel. The data is written to the DRAM, which is shared with the microcontroller. The microcontroller sends the data via Ethernet to the PC and can likewise receive control commands.

The PC is required for system control via user interface, complete signal processing and visualization. The latter two tasks are performed by the GPU (GTX1050TI, NVIDIA, Santa Clara, CA, USA). This enables parallel processing, and, consequently, higher frame rates.

In summary, the system is suitable for transmit beamforming, receive beamforming, and, thus, also for pulse echo operation. Next, the generation of an environmental image using several sequential pulse echo sequences is described.

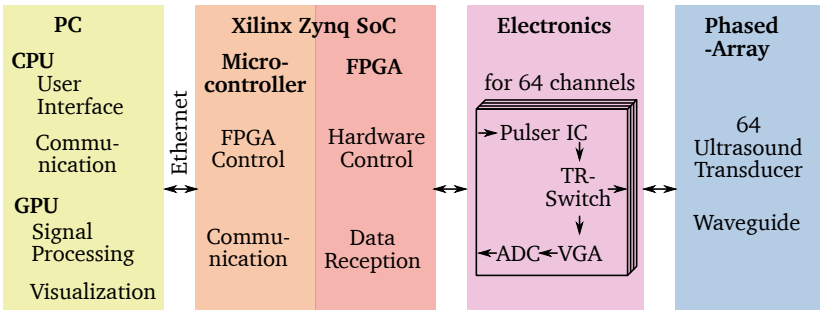

Figure 1. Phased-Array-System components and interfaces.

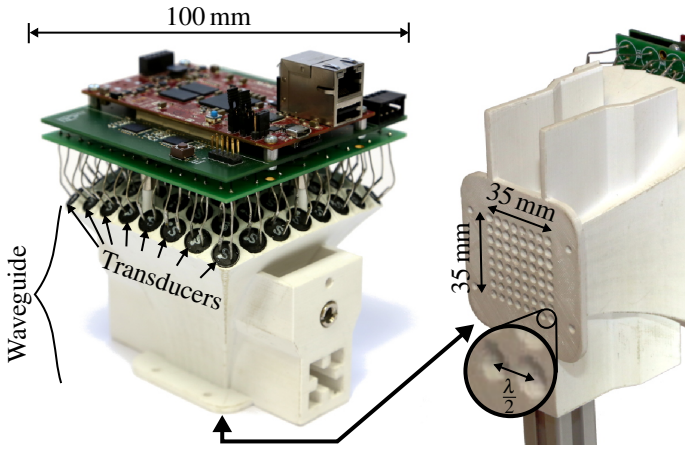

Figure 2. Phased-Array with waveguide and electronics attached. The waveguide reduces the inter-element spacing to $\lambda/2$ for grating lobe free beamforming. The analog electronics can switch between transmission and reception mode. The FPGA board controls the analog electronics and sends the acquired raw data to a PC via ethernet.

3 THEORY OF OPERATION

In order to generate a volume image of the space in front of the phased-array, the region of interest is discretized into raster beams. For each raster beam, a pulse echo sequence is performed. The recorded pulse echo sequence data of all transducers are processed and visualized in parallel. When all raster beams have been handled, the generated image represents the echo sources in the region of interest. The individual procedures are examined in detail in the following four sections.

3.1 Specification and discretization of the region of interest

The region of interest describes the section of space in front of the array which is to be represented in the generated image. The specification and discretization is defined once at the very beginning of the imaging process. All angles are specified for a spherical coordinate system [1], which simplifies uniform discretization [Figure 3(a)]. The region of interest is characterized in terms of range of view and field of view. The range of view R_{view} is determined by the time window T_{view} after the pulse transmission in which the transducer signals are recorded, i.e.

$$R_{\text{view}} = \frac{T_{\text{view}} \cdot \text{c}}{2}. \tag{1}$$

The field of view is determined with the minimum and maximum azimuth angles α_{min}, α_{max} and elevation angles ε_{min}, ε_{max}. By defining the angular step sizes α_{step} and $\varepsilon_{\text{step}}$, the field of view is discretized into $K \cdot G$ raster beams, where

$$K = \frac{\alpha_{\text{max}} - \alpha_{\text{min}}}{\alpha_{\text{step}}} + 1 \qquad \text{and} \qquad G = \frac{\varepsilon_{\text{max}} - \varepsilon_{\text{min}}}{\varepsilon_{\text{step}}} + 1 \tag{2}$$

are the number of azimuth and elevation angles, respectively [Figure 3(b)]. Each raster beam $b_{\text{k,g}}$ has a unique direction defined by two angles α_{k}, ε_{g}, i.e.

$$b_{\text{k,g}} = [\alpha_{\text{k}}, \varepsilon_{\text{g}}], \tag{3}$$

$$\alpha_{\text{k}} = \alpha_{\text{min}} + k \cdot \alpha_{\text{step}}, \qquad k = 0, \ldots, K-1, \tag{4}$$

$$\varepsilon_{\text{g}} = \varepsilon_{\text{min}} + g \cdot \varepsilon_{\text{step}}, \qquad g = 0, \ldots, G-1, \tag{5}$$

where k and g are the indices of the azimuth and elevation angles, respectively.

The following pulse echo sequence, signal processing and visualization are repeated for all $K \cdot G$ raster beams. A single raster beam is selected in each repetition.

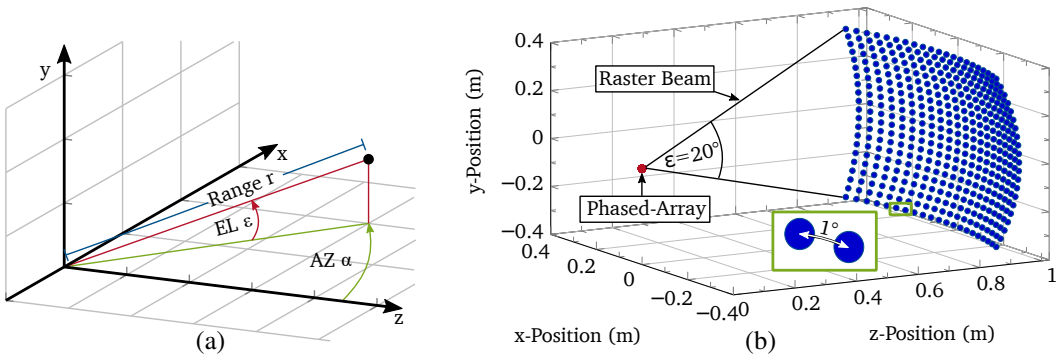

Figure 3. (a) Used spherical coordinate system. A point is defined with the range r, elevation ε and azimuth α angles. (b) Example region of interest raster with a field of view of $\alpha_{\text{min}} = \varepsilon_{\text{min}} = -10°$, $\alpha_{\text{max}} = \varepsilon_{\text{max}} = +10°$ and $\alpha_{\text{step}} = \varepsilon_{\text{step}} = 1°$. The range of view is $R_{\text{view}} = 1\,\text{m}$. A total of $21 \cdot 21 = 441$ raster beams are generated.

3.2 Pulse Echo Sequence

In the pulse echo sequence, a pulse is transmitted from the transceiver elements in the selected raster beam direction using far-field transmit beamforming. The resulting echo signals are recorded. All transducers are used for pulse transmission and echo reception. The sequence control for the transmission and reception is mainly performed by the FPGA and microcontroller. The specific tasks are explained in detail in the next paragraph.

First, the required time delays Δt_m of the transducer signals are calculated in order to send a pulse in the selected raster beam direction $b_{\mathrm{k,g}}$, i.e.

$$\Delta t_\mathrm{m} = \frac{1}{\mathrm{c}} \cdot \left[x_\mathrm{m} \sin(\alpha_\mathrm{k}) \cos(\varepsilon_\mathrm{g}) + y_\mathrm{m} \sin(\varepsilon_\mathrm{g}) \right] \tag{6}$$

where α_k, ε_g are the angles of the selected raster beam direction and x_m, y_m are the position of the m-th transducer [7]. A common offset is added to the time delays Δt_m to prevent negative, acausal values. Second, the individually time-delayed, unipolar 40 kHz square-wave burst signals $s_{\mathrm{tx,m}}$ are generated for all transducers. Each burst has a length of 40 periods and an amplitude of 20 $\mathrm{V_{pp}}$. After pulse transmission, all TR-switches are switched to receive mode. Finally, the digitized, received signals of the transducers $s_{\mathrm{rx,m}}$ are read in parallel in the time window T_view and sent to the PC.

3.3 Signal Processing

The PC receives the raw data of all transducers $s_{\mathrm{rx,m}}$ and buffers them into the GPU memory. The complete signal processing and visualization pipeline is performed by the GPU in parallel using OpenGL Compute Shaders. The signal processing pipeline consists of parallel preprocessing blocks, the Delay-and-Sum beamformer (DAS) and the envelope generator (Figure 4). All signal processing is performed in the time domain.

The bandpass filter is designed as an FIR filter with 176 filter taps, a center frequency of 40 kHz and a bandwidth of 2 kHz using the window method [8]. The impulse response of the filter is convoluted with the input signals.

Next, the signals are upsampled and interpolated by a factor of 25 using zero-stuffing and sinc-filtering [8]. This interpolation step is required to reduce the smallest possible time delay for the Delay-and-Sum beamformer, which improves the directional step size of the spatial filter.

The Delay-and-Sum beamformer delays the preprocessed signals according to (Equation 6) and then sums them up into a single signal. The added time delays ensure that the echo pulse captured from the selected raster beam direction is in-phase for all transducer signals. This results in a maximum echo amplitude during summation. At the same time, interference echoes originating from other directions are suppressed. The Delay-and-Sum beamformer therefore functions as a spatial filter [1].

Finally, the envelope $s_\mathrm{env}(n)$ of the summed signal is determined. For this, the analytic signal $s_{\mathrm{sum,a}}(n)$ is computed and the complex absolute value is calculated, i.e.

$$s_{\mathrm{sum,a}}(n) = s_\mathrm{sum}(n) + \mathrm{j} \cdot s_\mathrm{sum}(n) * h(n), \tag{7}$$

$$s_\mathrm{env}(n) = \left| s_{\mathrm{sum,a}}(n) \right|, \tag{8}$$

where n is the sample index, j is the imaginary unit and $h(n)$ is the time-discrete Hilbert transform [8]. The computed envelope signal represents the echo amplitude from the selected raster beam direction. It is the basic component of the following visualization.

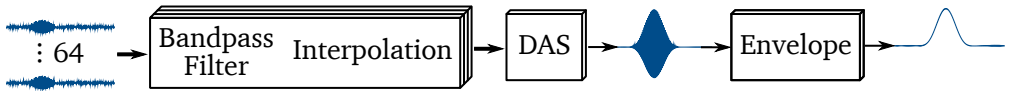

Figure 4. Signal processing pipeline for generating a summed, spatial filtered echo signal using 64 raw transducer signals. The FIR bandpass filtering and interpolation are performed in parallel for all transducer signals.

3.4 Visualization

In the visualization process, the horizontal sectional image (B-Scan) and the three-dimensional image (3D C-Scan) of the region of interest are built up step by step with each new envelope (Figure 5).

First, the envelope is color-coded based on its amplitude and flattened. Next, it is positioned according to the selected raster beam direction. In combination with the previously generated envelopes, a grid view is created. After the scan conversion from spherical to cartesian coordinates, the grid view shows the correct positions of the envelopes with respect to the real space. Finally, the gaps of the grid view are filled using linear image interpolation.

The same steps are used to generate the 3D C-Scan. However, both the horizontal and the vertical angle are considered for the positioning of the envelope. Additionally, alpha blending is used for image interpolation to make only the high echo amplitudes visible. In the next section, the implemented imaging method is characterized.

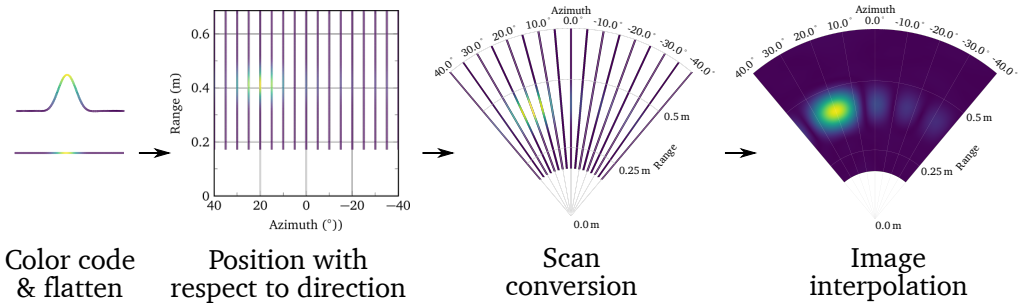

Figure 5. Visualization pipeline for generating a B-Scan. The envelope is color coded, flattened and positioned according to its originating direction. After the scan conversion from spherical to cartesian coordinates, the grid view represents the correct envelope positions. The grid view is linearly interpolated.

4 MEASUREMENT RESULTS

The phased-array and the imaging method are evaluated in an anechoic chamber with focus on the transmit and receive characteristics, the lateral and axial resolution as well as the blind zone and maximum range.

4.1 Measurement setup

In the anechoic chamber the phased-array is mounted on two rotational axes at the end of a 6 m long rail (Figure 6). On the rail, there is a movable slide on which different targets can be attached. Depending on the measurement, an ultrasonic transducer, a calibrated microphone or two hollow steel spheres with a diameter of 100 mm are used as targets. Using the rotational axes and the slide, the targets can be positioned freely in the coordinate system of the phased-array. In order to make the measurement results comparable, the VGA amplifies the recorded signals with a constant amplification factor of 46 dB in all measurements.

4.2 Transmit and receive characteristics

The transmit and receive characteristics provide information about the spatial sound distribution during transmit beamforming and the spatial filter response of the Delay-and-Sum beamformer, respectively. Significant parameters are the -3 dB main lobe width and the side lobe level. The main lobe widths of both characteristics define the separability of laterally adjacent echo sources, i.e. the lateral resolution. The side lobe levels determine the relative amplitude of the resulting phantom reflections for a present echo source. The phantom reflections degrade the overall image quality.

For measuring the transmit characteristic, the calibrated microphone is used as target and positioned at a distance

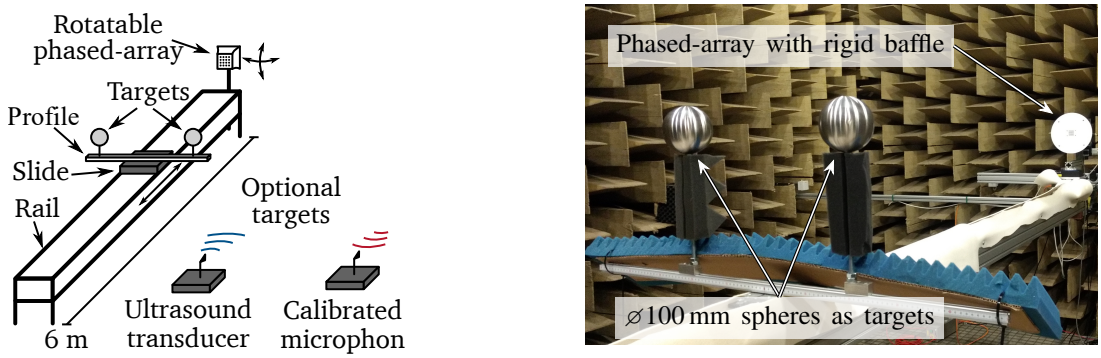

Figure 6. Measurement setup in the anechoic chamber. The phased-array can be rotated mechanically with two DOF using rotational axes. For the transmit characteristic measurement, the target is replaced with a calibrated microphone. For measuring the receive characteristic, the target is replaced with an ultrasound transducer. Otherwise two spheres with a diameter of 100 mm are used.

of 2 m. The phased-array sequentially sends pulses to the direction (EL, AZ) = $(0°, 0°)$ in the array coordinate system. The microphone records the incoming sound pressure. After each transmission, the array is rotated mechanically in 1° steps. The measurement is repeated until all angles of the field of view in Figure 7(a) have been selected.

For the receive characteristic measurement, an ultrasonic transmitter is used as target and positioned at a distance of 2 m. The transmitter sends a pulse to the phased-array, which is in receive-only mode. All ultrasound transducers pick up this pulse and generate 64 signals. With this data set, the spatial Delay-and-Sum filter sequentially generates the summed signal envelopes for all directions of the field of view in Figure 7(c).

The measurement results show an asymmetry in both characteristics. This is explained by the fact that the properties of the individual transducers differ. First, there are deviations in the radiated signal power and the receiving sensitivity. Second, there are variations in the resonant frequencies and phase responses. The 3 dB mainlobe widths are 21° and 12°, the maximum sidelobe level is −7.2 dB and −11.5 dB for the transmit and receive characterisic, respectively. Next, the imaging properties are examined in a pulse echo scenario.

4.3 Axial and lateral resolution, blind zone and maximum range

In order to measure the axial and lateral resolution, two spheres are placed either axially or laterally next to each other on a profile mounted on the slide. The distance d between the spheres is adjustable. B-scan images are used to locate the two targets. The targets are defined as separable if there are two distinct local maxima. The minimum distance d_{min} at which the spheres are seperable determines the axial resolution $d_{\text{res, ax}}$ or the lateral resolution $\alpha_{\text{res, lat}}$ [9], i.e.

$$d_{\text{res, ax}} = d_{\text{min, ax}}, \qquad \alpha_{\text{res, lat}} = 2 \cdot \arctan\left(\frac{d_{\text{min, lat}}}{2 \cdot r}\right). \tag{9}$$

The latter is specified as an angle, where r is the range between the phased-array and the profile. For the measurement, the profile is positioned at a range of $r = 3$ m and direction AZ 0°, EL 0° (Figure 8).

The axial resolution is $d_{\text{res, ax}} = 20$ mm using a pulse length of 40 periods. If the length of the transmitted pulse is reduced to 10 periods, the axial resolution improves to $d_{\text{res, ax}} = 15$ mm. However, due to the weaker transmission amplitude when a shorter pulse is transmitted, the maximum detection range is reduced.

The lateral resolution is $\alpha_{\text{res, lat}} = 12°$. At a range of $r = 3$ m, this corresponds to a minimum distance between the spheres of $d_{\text{min, lat}} = 0.63$ m. As the range r to the phased-array increases, the required distance $d_{\text{min, lat}}$

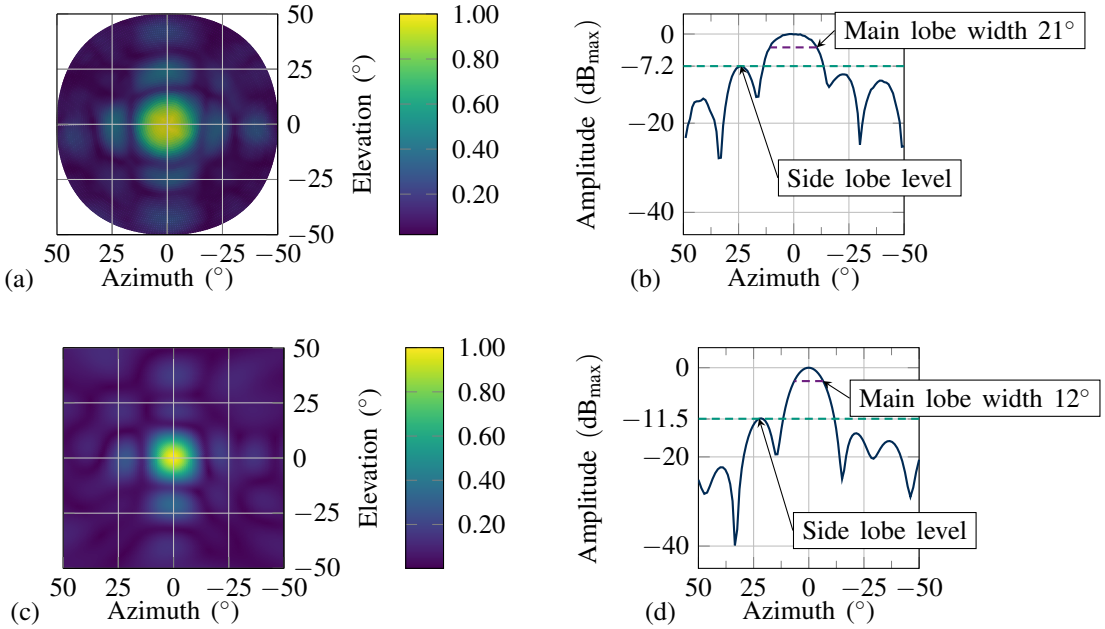

Figure 7. (a) Measured transmit characteristic. (b) Horizontal section of a) at EL 0°. (c) Measured receive characteristic. (d) Horizontal section of c) at EL 0°.

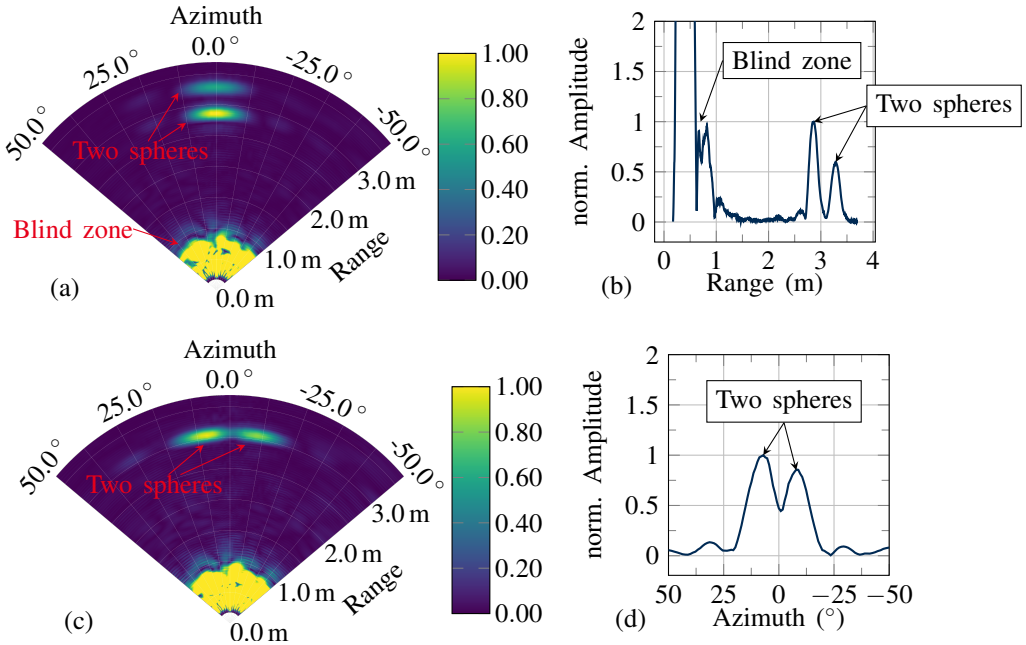

Figure 8. (a) Axial resolution B-Scan of two 100 mm spheres located at AZ 0°, EL 0°, 2.8 m and 3.2 m, respectively. (b) Axial section of the B-Scan in a) at AZ 0°. (c) Lateral resolution B-Scan of two 100 mm spheres located at 3 m, EL 0°, AZ −7° and AZ 7°, respectively. (d) Lateral section of the B-Scan in c) at 3.1 m.

between the spheres for separability increases as well according to (Equation 9).

In the B-Scan images shown, there is a blind zone of up to 1 m. In this zone, the transmitted pulse and the oscillation of the transducers masks incoming echoes. With a suitable amplitude scaling of the color bar for close targets, echoes from a range of 0.5 m and above can still be detected. The blind zone defines the lower limit of the echo detection.

The upper limit is determined by the attenuation of the medium, since amplitudes from distant echos are too weak to be detected. The maximum detection distance is validated up to 6 m, which corresponds to the maximum length of the rail.

5 CONCLUSION

The air-coupled ultrasound imaging system presented can generate 2D B-Scans and 3D C-Scans in a horizontal and vertical field of view of $\pm 50^\circ$ and a range of 0.5 up to 6 m. The system consists of custom electronics with 64 channels, an FPGA and microcontroller, as well as a PC. The GPU accelerated parallel signal processing and visualization enables fast processing times.

The frame rates for a B-Scan (2 FPS) and a 3D C-Scan (0.1 FPS) are mainly limited by the sound propagation time. For optimization of the frame rate, a single transducer can be used to transmit a spherical wave. This way, echoes from all spatial directions are generated simultaneously with a single pulse.

Another main drawback is the lateral resolution of 12°. To improve this, advanced direction of arrival estimators will be evaluated in the future, for example techniques based on compressed sensing [10].

REFERENCES

[1] Johnson, D. H.; Dudgeon, D. E. Array Signal Processing: Concepts and Techniques, 1st ed.; Prentice Hall: New Jersey, 1993. pp. 90-91, 9-10, 112-128.

[2] Jäger, A. et. al., Air-Coupled 40-KHz Ultrasonic 2D-Phased Array Based on a 3D-Printed Waveguide Structure. In IEEE International Ultrasonics Symposium (IUS); 2017.

[3] Langen, A. Ein Verfahren zur Konstruktion anwendungsoptimierter Ultraschallsensoren auf der Basis von Schallkanälen, 1st ed.; Springer: Berlin-Heidelberg, 1993.

[4] Takahashi, T.; Takahashi, R.; Jeong, S. Ultrasonic Phased Array Sensor for an Electric Travel Aids for Visually Impaired People. In Proc. SPIE 6794, ICMIT: Mechatronics, MEMS, and Smart Materials; 2007.

[5] Unger, A. et. al. Versatile Air-Coupled Phased Array Transducer for Sensor Applications. In IEEE SENSORS; 2015.

[6] Rutsch, M. et. al. Extending the Receive Performance of Phased Ultrasonic Transducer Arrays in Air down to 40 KHz and Below. In IEEE International Ultrasonics Symposium (IUS); 2015.

[7] Schmerr Jr., L. W. Fundamentals of Ultrasonic Phased Arrays, 1st ed.; Springer: New York, 2015. pp. 169-172.

[8] Oppenheim, A. V.; Schafer, R. W.; Buck, J. R. Discrete-Time Signal Processing, 2nd ed.; Prentice Hall: New Jersey, 1999. pp. 465-474, 150-153, 775-801.

[9] Hoskins, P.; Martin, K.; Thrush, A. Diagnostic Ultrasound: Physics and Equipment, 2nd ed.; Cambridge University Press: Cambridge, 2010. pp. 64-66

[10] Yang, Y.; Pesavento, M.; Chatzinotas, S.; Ottersten, B. Successive Convex Approximation Algorithms for Sparse Signal Estimation with Nonconvex Regularizations. In IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM); 2018.