Development of an airborne circular sonar for research and teaching

Blair Bonnett[1], Kaushikk V. N.[1], Sudhanshu Apte[1], and Thomas Fickenscher[1]

[1] Chair for Radio Frequency Engineering, Helmut Schmidt University, 22043 Hamburg, Germany

Corresponding author: *Blair Bonnett, Lehrstuhl für Hochfrequenztechnik, Fakultät für Elektrotechnik, Helmut-Schmidt-Universität, 22039 Hamburg, Germany; blair.bonnett@hsu-hh.de*

Abstract: *Sonar is largely considered to be an underwater technique due to the predominance of other measurement methods for above-water applications. However, in-air echolocation is used by various animals (bats being the obvious example). It has also been utilised in some electronic sensors, for example basic range finders and fluid level sensors. Recent publications have shown that synthetic aperture sonar (SAS) images can be formed from data collected with standard audio speakers and microphones. This allows complex scenes to be imaged in a controlled laboratory environment for validating processing and analysis techniques, although differences in the acoustic response of the targets in air versus in water must be considered. Such a system can also be a valuable tool for teaching and demonstrating sonar principles.*

This paper introduces an airborne circular SAS system under development. It uses a Peerless OC16SC00-04 speaker (16 mm diameter) and Syntiant SPG08P4HM4H-1 MEMs microphones (4 mm by 2 mm). An AMD Zynq system-on-chip comprising a dual-core ARM Cortex-A9 processor and AMD 7 series FPGA is used to control the system and synchronise the components. The system is operated over a standard Ethernet connection. The small size of the microphones and the ability of the FPGA to handle many microphones simultaneously (62 in the current setup) allows for receiver arrays analogous to those used by underwater systems. This creates an opportunity to implement array-based processing techniques such as interferometry or motion compensation.

As with underwater sonar, precise calibration of the system and knowledge of the environment (in particular, the speed of sound) are required to obtain good results. We show the current status of the system and its calibration along with some applications. Future development plans, including for alternative hardware designs, will be outlined. All custom hardware designs and software for using the system will be released under open source licenses.

Keywords: *synthetic aperture sonar, airborne sonar, sonar teaching*

1. INTRODUCTION

Sonar is largely considered to be an underwater technique. Echolocation was first investigated for obstacle detection in the wake of the *Titanic* sinking and was then developed (initially under the name ASDIC) for anti-submarine detection. Subsequent research and usage have mostly focused on underwater applications. Although other methods (optical, radar, lidar etc.) are more prevalent for above-water applications, the principles of sonar also work in air. In nature, many species of bat use echolocation for navigation and hunting as do some other mammals and birds, while some prey species have evolved countermeasures [1]. In-air sonars have been used for a variety of applications such as aids for blind people [2] and to develop and test general sonar processing techniques [3] while ultrasonic distance sensors are commonly used in robotics and fluid level monitoring.

Recent publications have shown that synthetic aperture sonar (SAS) processing can be used to form images from data collected with standard audio speakers and microphones [4, 5]. This can be used to image complex scenes in controlled environments to generate datasets for validating processing and analysis techniques [6]. Although the acoustic response of specific targets will differ in air compared to water, they exhibit the same range of complex scattering effects (occlusion, multiple scattering, elastic scattering etc) that is observed in underwater sonar.

In this paper, we will introduce an airborne circular SAS we are developing. This is based around micro-electromechanical systems (MEMS) microphones which are sufficiently small to form receiver arrays, similar to those commonly used in underwater SAS systems. We start by describing the electronics used in the system and the platform it is mounted to. This is followed by details of the firmware used to control the system and capture the microphone outputs, and the subsequent processing to form SAS images from this data. Some example images captured by the airborne sonar are shown before we conclude with plans for improving and extending this system.

2. ELECTRONICS

A MEMS microphone uses a thin membrane suspended above a fixed backplate as the sensor. The surrounding case is sealed apart from one hole known as the audio port. When a pressure wave enters the audio port, it deflects the membrane. This change in distance between the backplate and the membrane can then be detected as a change in capacitance. They are small and widely used in consumer devices such as smartphones, wearable electronics and televisions.

For this project, the Syntiant (formerly Knowles) SPG08P4HM4H-1 MEMS microphone was selected. This comes in a 4 mm × 2 mm package with an acoustic port diameter of 0.65 mm, an overload point[1] of 120 dB_{SPL}, an SNR of 64 dB with a 1 Pa pressure wave and a total harmonic distortion of less than 0.1 % for the same 1 Pa input. Although designed for use in the audible range, the ultrasonic response plot in the datasheet suggests a −3 dB cutoff at approximately 44 kHz. The output of the microphone is a single-bit pulse-density modulation (PDM) signal (i.e., the output of a sigma-delta modulator) referenced to a clock signal of up to 4.8 MHz supplied to the microphone. Each microphone can be configured to output its data on either the rising or falling edge of the clock signal, allowing two microphones to share the same data line (often utilised for stereo pairs in consumer devices).

[1] Note that all sound pressure levels mentioned in this paper are relative to the 20 µPa reference level used in airborne audio (the limit of human hearing) instead of the 1 µPa reference normally used in underwater acoustics.

A printed circuit board (PCB) was designed to hold eight of these microphones with a spacing of 8.5 mm (approximately half a wavelength at a frequency of 20 kHz). A 12-wire ribbon cable provides power and the sampling clock, and returns the eight data lines. Solder bridges are used to configure whether the microphones output on the rising or falling edge of the clock. To avoid excess load on the input clock signal, a Texas Instruments LMK1C1108 clock buffer is used to distribute a copy of it to each microphone.

A Peerless OC16SC00-04 tweeter is used as the transmitter. This has a circular face with a diameter of 16 mm, a sensitivity of 81.8 dB_{SPL} at 1 m for a 1 W input signal and a power rating of 100 W (which should then correspond to 87.8 dB_{SPL} at 1 m). Its nominal frequency range is 2.5 kHz to 20 kHz; the datasheet indicates good on-axis performance up to 40 kHz although the off-axis strength decreases before this. An STMicro TDA7498E class D audio amplifier in a manufacturer-supplied evaluation kit rated for 160 W is used to drive the speaker.

The target to be imaged is mounted on a rotating stage (technically, this forms an inverse circular SAS as the target moves while the imaging system is stationary). A Phidgets RB-Phi-267 stepper motor is used to turn the stage. The motor has a step angle of 1.8° and is connected to a 1:99.51 planetary gearbox for a final step angle of 0.0181°. The Pololu Tic 36v4 stepper motor driver is used to control the motor.

To control the system, a Digilent Zybo Z7-20 development board is used. This contains an AMD (formerly Xilinx) Zynq-7020 system-on-chip (SOC) which integrates a dual-core ARM Cortex-A9 processor and an FPGA. This combination allows custom hardware interfaces to be designed in a hardware description language and connected to the processor via standard communications buses. The development board contains a number of peripherals, of which the gigabit Ethernet connector, audio codec and expansion connectors are used in this project. Two PCBs have been built to interface to this development board: one for communication with the motor controller and another with connectors for up to eight of the microphone boards. Six of these connectors can handle all eight microphones from a PCB while the final pair have one connection unavailable as the corresponding output from the development board is needed for the clock signal. This allows up to 62 microphones to be connected to the system.

3. PLATFORM

The system (shown in Figure 1a) is mounted onto a 1.2 m long by 0.6 m wide wheeled trolley so that it can be easily moved for teaching, demonstrations etc. The rotating stage consists of an 80 cm diameter disc of 6 mm thick perspex with the motor mounted from underneath. A series of captive ball bearings provide support to the edge of the stage. For some experiments, a 10 mm thick and 20 mm high wooden border is mounted around the perimeter of the stage to contain loose material; Figure 1a shows the stage filled with gravel. The PCBs are mounted at the other end of the trolley from the rotating stage while a lower shelf holds the power supplies.

The speaker and microphones are mounted in a 3D-printed holder allowing their arrangement to be easily customised for a given scenario. In the results presented here, an interferometric pair of 16-element receiver arrays with a baseline of 51 mm is used with the transmitter located equidistantly between and aligned with the centre of the arrays. This setup is shown in Figure 1b. The centre of the speaker is located 750 mm laterally from the centre of the rotating stage. The head is tilted down with the height of the transmitter set so that its boresight is pointing at the centre of the stage, i.e., the height depends on the tilt angle.

(a) *(b)*

Figure 1: The airborne sonar system: (a) rotating stage, sound head and PCBs (power supplies are on a shelf below the main table) and (b) a close-up of the sound head. The speaker is in the centre, and two horizontal linear arrays of MEMS microphones (each made up of two of the microphone PCBs) can be seen, one above and one below the speaker.

4. FIRMWARE

The VHDL language is used to implement the firmware for the FPGA side of the Zynq-7020. The main sonar peripheral has a number of responsibilities:

- Configure the audio codec to act as a digital-to-analogue converter to drive the amplifier.
- Load signals to transmit from the processor.
- Handle the signals returned from all microphones, including separating the signals from the rising and falling edges of each input, and allowing the processor to specify reordering of the input channels for flexibility.
- When requested, perform a ping. This involves sending the samples of the previously loaded signal to the audio codec for transmission, synchronising the receiver to the transmitter, and recording a set number of samples from the microphones.

The recording utilises direct memory access to directly place the samples in RAM without any processor overhead. As this is performed by the FPGA, all the microphones inputs are captured simultaneously; the accuracy in relative timing between different microphones, and between the speaker and the microphones, is determined by the electronics.

Currently only a single core of the processor is used. This is programmed in C utilising the support libraries provided by the vendor along with a number of custom libraries. It is designed to act as a simple server over the Ethernet connection provided on the development board. A client can connect to the server, configure the sonar and then start a capture. The recorded samples are returned to the client formatted in a simple packet.

5. PROCESSING

The processing is performed in the Python programming language with a number of common libraries, namely NumPy, SciPy and Xarray. The processing code communicates with the sonar over the network interface described previously. As recorded samples are received, they are placed in a structured dataset and subsequently stored using the netCDF binary format.

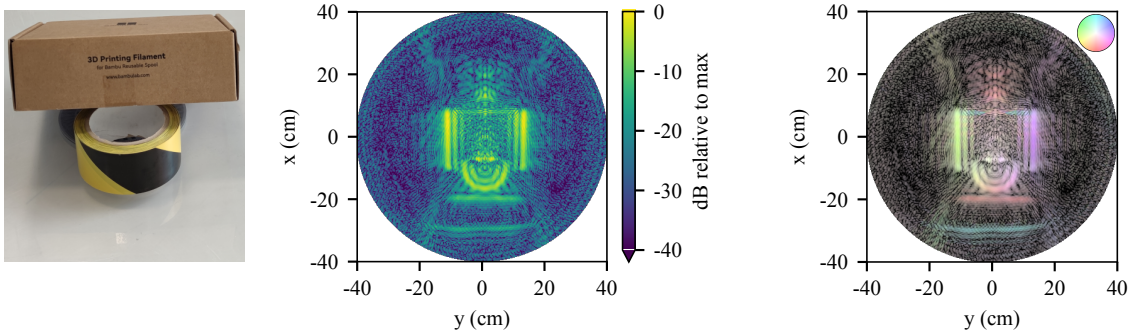
As mentioned in the electronics section, the output of the microphones is a single-bit PDM signal; we use the maximum 4.8 MHz sampling rate. This needs filtering and downsampling to provide a higher-resolution signal at lower data rate. When performed in real-time, this requires careful design of the filters to ensure both computational efficiency and sufficient attenuation of the high-frequency noise in the original signal. In this case, we have no real-time requirements and so standard finite impulse response (FIR) filters can be used. SciPy's `signal.decimate` function is used to downsample the signal in two stages, first by a factor of eight after applying a low-pass FIR filter with 160 taps and then by a factor of six after applying another low-pass FIR filter with 120 taps. In both stages the group delay of the filters is compensated for to avoid introducing phase errors. The resulting signal is sampled at 100 kHz.

The microphone outputs have a low-frequency drift. To remove this and any other out-of-band noise, a FIR bandpass filter with transition frequencies 5 kHz either side of the transmitted spectrum is applied. Pulse-compression is then performed by correlating the received echo with a replica of the transmitted signal. This correlation is performed in the Fourier domain with a guard band to avoid circular effects. The pulse-compressed signal is converted to its complex analytic representation and a range-varying gain applied before delay-and-sum backprojection is used to form an image.

6. EXAMPLE IMAGES

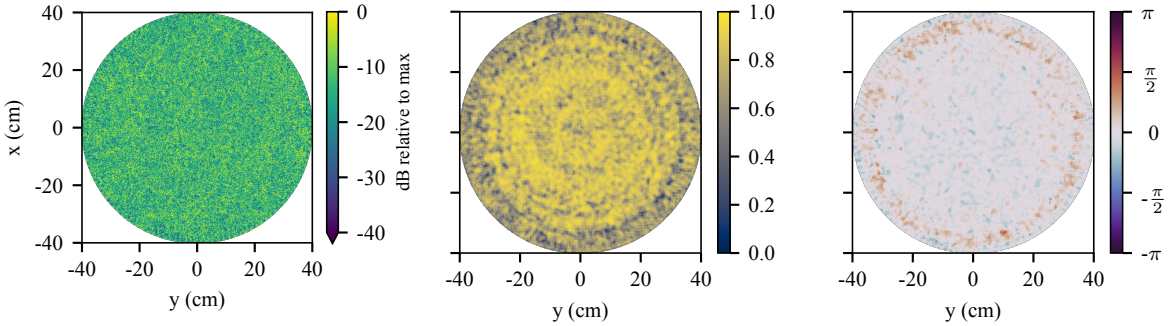
In initial testing, it was observed that the received energy began to decrease as the frequency exceeded 20 kHz. As both the speaker and microphones should be able to operate up to 40 kHz, it is assumed that the amplifier is the cause of this limitation. To accommodate this, a 10 kHz to 30 kHz hyperbolic frequency-modulated chirp of 2 ms duration was used in the data collection for the images presented here. This modulation ensures a higher proportion of the chirp uses the lower part of the spectrum. The stage was rotated 3° between pings which results in an approximately 50 % overlap of the phase centre arrays from ping to ping. The images presented here have been reconstructed at a 2 mm resolution with areas outside the rotating stage removed.

The first scene is a cardboard box (205 mm × 220 mm × 70 mm) tilted up on a roll of tape (108 mm diameter, 50 mm high), shown in Figure 2a. The SAS image from the upper array (Figure 2b) shows that the edges of the targets reflect strongly. Some multipath echoes can be observed as can some resonance patterns. The background is mostly noise as the rotating stage is highly specular. The repeating pattern around the outer edge of the circle is due to a stationary reflector outside the stage which is smeared around the entire circle. Figure 2c shows a colour-by-aspect (CBA) version of the reconstructed image. Following the approach of [7], this encodes the direction from which each pixel appears brightest as the hue, the inverse entropy of the pixels as the saturation and the intensity as the value. This means pixels which were observed from all directions tend to greyscale while pixels which were observed primarily from one direction have the corresponding colour. The embedded colour wheel shows the mapping of colour to view direction. From this, it can be seen that, as expected, the sides and edges of the targets were mostly imaged from the corresponding side of the stage.



(a) Scene *(b) SAS image from upper array* *(c) CBA image from upper array*

Figure 2: Results from imaging a box tilted up on a roll of tape with a 30° grazing angle.



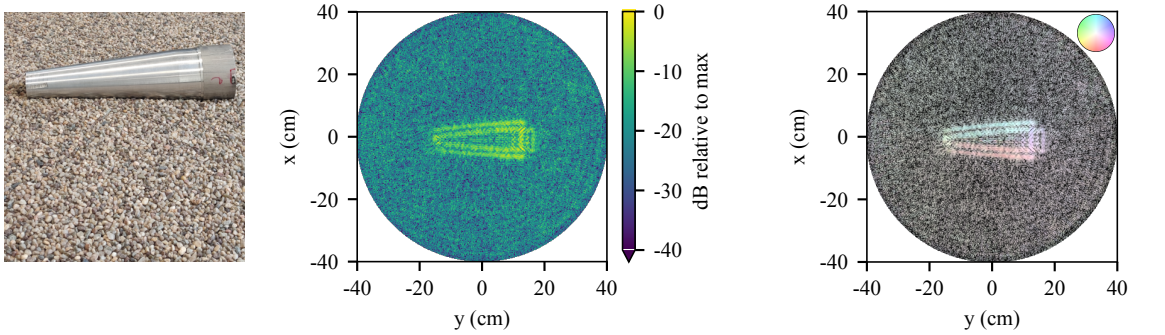
(a) SAS image from upper array *(b) Coherence between arrays* *(c) Interferometric phase*

Figure 3: Results from imaging a bed of gravel with a 30° grazing angle.

For the remaining results, the rotating stage was filled with gravel to a depth of 20 mm. The nominal size range of the gravel is 3 mm to 5 mm although both larger and smaller pieces are present in small quantities. Figure 3a shows the SAS image reconstructed from the upper array with no extra target present. This has the speckle nature expected of a rough background. The coherence between images from the two arrays (Figure 3b) is high in the middle of the stage but decreases near the edges. It is hypothesised that the edges are not insonified on every ping (i.e., are outside the full-view area) at the high end of the spectrum due to the narrow beamwidth of the transmitter. These areas of the image would then have a lower SNR leading to reduced coherence. The interferometric phase between the images (Figure 3c) is low in the centre of the stage, corresponding to a surface with no significant height variations. The higher values around the edge correspond to areas of low coherence and as such can be ignored.

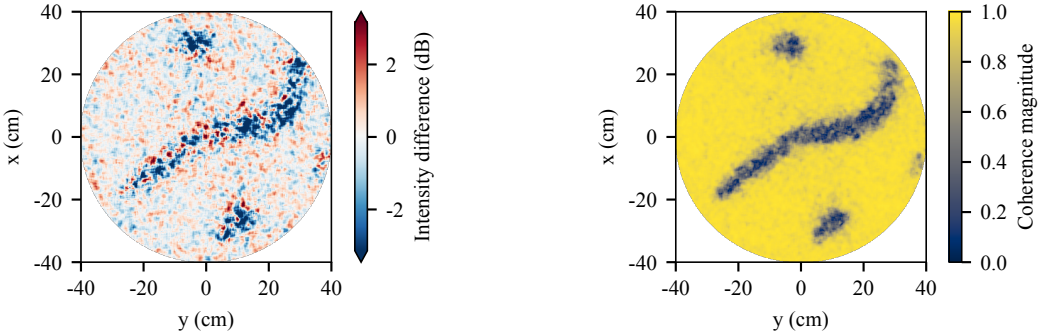
A shell-like shape machined from aluminium was then placed on the gravel. This is 290 mm long and tapers from 80 mm to 40 mm in diameter with an untapered portion at the large end. This scene is shown in Figure 4a and the SAS image from the upper array in Figure 4b. A double echo can be seen; it is not known whether this is a multipath reflection or an elastic scattering effect. The CBA image of Figure 4c shows that the echoes from the target were observed directionally while the speckle pattern from the gravel was captured from most angles.

Finally, the suitability of the system for change detection was investigated. Two images of the gravel were captured. Between these captures, some trenches were created in the gravel with a spade. As shown in Figure 5, both incoherent and coherent change detection can be successfully performed on these images.



(a) Scene *(b) SAS image from upper array* *(c) CBA image from upper array*

Figure 4: Shell-like target imaged at a 15° grazing angle.



(a) Incoherent change detection *(b) Coherent change detection*

Figure 5: Change detection between two images of a bed of gravel. A 7 by 7 pixel window (1.4 cm by 1.4cm) was used to smooth the results in the incoherent case. The same size window was used when estimating the coherence.

7. CONCLUSION

The initial results presented here show that MEMS microphones are suitable for collecting airborne sonar data which can be coherently processed to form SAS images. Furthermore, their small size allows them to be formed into arrays which offers the potential for testing standard sonar algorithms which require such arrays, for example bathymetric estimation from interferometry. The successful application of change detection suggests that the system produces repeatable results (the only significant changes were those deliberately introduced to the scene).

There are improvements to the current system we wish to implement. As mentioned above, the transmitter amplifier appears to be limiting the maximum frequency to a point below that which the speaker and microphones should be capable of. Both the digital-to-analogue conversion and the amplifier also introduce delays between a ping being requested and the chirp starting. Although this can be calibrated to a point where any receiver timing error due to these delays do not cause significant problems to image formation, it would be preferable to eliminate them. A class D or similar amplifier driven directly from the FPGA would allow full control over the transmitter and could be designed to work with higher frequencies. Full calibration of the transmitter source levels, receiver sensitivity and beampatterns of both transmitter and receiver is also planned.

The images presented here were processed assuming a sound speed of 343 m/s. Some work has been performed towards measuring the sound speed with the system. In the future, we intend to add temperature sensors to the system to also allow standard temperature-based sound speed models to be used. Pressure and humidity sensors can also be added. These parameters typically have minimal impact on the sound speed, but may be more significant for modelling the attenuation of the sound.

A linear variant of this system is planned in the future. The ability to add non-ideal motion, for example yaw or sway, to the capture geometry would allow the collection of data more representative of an underwater system. Interfacing a MEMS receiver array to a single-board computer such as a Raspberry Pi instead of requiring the use of a CPU + FPGA system would make the system more accessible for others to implement, albeit with capacity for fewer microphones and potentially with less timing accuracy.

The designs for this system have been released under open-source licenses and can be found at `https://github.com/airsonar`. The circuit designs are available under the CERN Open Hardware Licence and source for the firmware under the BSD 2-clause plus patent license.

ACKNOWLEDGEMENTS

The authors would like to thank the technical staff of Helmut Schmidt University for their assistance in building the airborne sonar and the experimental targets. We would also like to thank Dr. Holger Schmaljohann of WTD 71 for valuable feedback and suggestions regarding this project.

REFERENCES

[1] W. E. Conner and A. J. Corcoran. “Sound Strategies: The 65-Million-Year-Old Battle Between Bats and Insects”. *Annual Review of Entomology*, vol. 57, 2012, pp. 21–39.

[2] L. Kay. “A sonar aid to enhance spatial perception of the blind: engineering design and evaluation”. *Radio and Electronic Engineer*, vol. 44, no. 11, 1974, pp. 605–627.

[3] P. T. Gough, D. de Roos, and M. J. Cusdin. “Continuous transmission FM sonar with one octave bandwidth and no blind time”. *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 131, no. 3, 1984, pp. 270–274.

[4] T. E. Blanford, J. D. McKay, D. C. Brown, J. D. Park, and S. F. Johnson. “Development of an in-air circular synthetic aperture sonar system as an educational tool”. *Proceedings of Meetings on Acoustics*, vol. 36, no. 1, Aug. 2019, p. 070002.

[5] T. E. Blanford, L. Garrett, J. D. Park, and D. C. Brown. “Leveraging audio hardware for underwater acoustics experiments”. *Proceedings of Meetings on Acoustics*, vol. 46, no. 1, Nov. 2022, p. 030002.

[6] T. E. Blanford, D. P. Williams, J. D. Park, B. T. Reinhardt, K. S. Dalton, S. F. Johnson, and D. C. Brown. “An in-air synthetic aperture sonar dataset of target scattering in environments of varying complexity”. *Scientific Data*, vol. 11, no. 1, Nov. 2024, p. 1196.

[7] D. S. Plotnick and T. M. Marston. “Utilization of Aspect Angle Information in Synthetic Aperture Images”. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, 2018, pp. 5424–5432.