

# Low-cost One-bit MEMS Microphone Arrays for In-air Acoustic Imaging Using FPGA's

Robin Kerstens, Dennis Laurijssen  
University of Antwerp  
Faculty of Applied Engineering - CoSys-Lab  
Groenenborgerlaan 171, Antwerpen  
robin.kerstens@uantwerpen.be, dennis.laurijssen@uantwerpen.be

Jan Steckel  
University of Antwerp & Flanders Make  
Faculty of Applied Engineering - CoSys-Lab  
Groenenborgerlaan 171, Antwerpen  
jan.steckel@uantwerpen.be

**Abstract**—Recent advancements in MEMS microphone integration have led to the production of low-cost digital MEMS microphones which are suitable for applications in the ultrasonic acoustic spectrum. Using these microphones instead of their analog counterparts can greatly simplify board- and system-design and reduce overall construction cost of an array built with these sensors. In this paper we will propose an architecture for using these low-cost one-bit MEMS microphones to construct a microphone array and demonstrate their applicability in beamforming-based acoustic imaging algorithms targeted at 3D sonar sensors. We developed a small prototype array to demonstrate that these microphones are well suited for array applications, and the proposed system architecture, developed around an FPGA-SOC, is shown to be ideally suited to interface with a large number of these microphones.

## I. INTRODUCTION

As has been shown previously, combining multiple microphones to form array based sonar sensors provides a great platform for applications such as beamforming or acoustic imaging [1]. Steering the array using signal processing techniques also has advantages over mechanically steering a directional microphone in a certain direction. Indeed, the lack of moving parts, the accompanying precision errors, and the wear and tear that goes along with such moving systems, make a static system more favorable. Theoretically it is simple to increase the spatial resolution of a microphone array, for example by adding microphones to the array. However, as we have argued before [1], this also leads to an increase in component count and board-design complexity, as each microphone needs an analog frontent and an analog-to-digital converter (ADC).

To counter these drawbacks companies like Knowles [2] have succeeded in making digital MEMS microphones with a built-in sigma-delta ADC [3] with an acoustic bandwidth of up to 100kHz, so that the needed additional circuitry is minimal and all other interfacing can use a simple 1-bit binary interface such as normal General Purpose Input Output (GPIO). An array constructed using these microphones would greatly facilitate the construction of medium-sized array sensors (consisting of 20-100 elements), which we have shown to be suitable for solving complex tasks like beamforming and 3D acoustic imaging. Tasks like these greatly depend on the quality of the microphones used, with characteristics

such as phase consistency between different units in a system and linearity as key characteristics. These issues will be addressed in this paper, where the microphones will be used for a beamforming application. Furthermore, we will present a hardware architecture constructed around an FPGA System-On-Chip which combines a dual-core ARM processor running at high clock speeds with a medium-sized FPGA. This hardware architecture is ideally suited for massively parallel data-acquisition and high-level signal processing.

The rest of the paper is structured as follows. In section II an explanation of the acoustic imaging techniques in this paper will be given. In section III an overview of the used hardware platform and experimental setup will be provided. The results are addressed in section IV and finally the conclusion in section V.

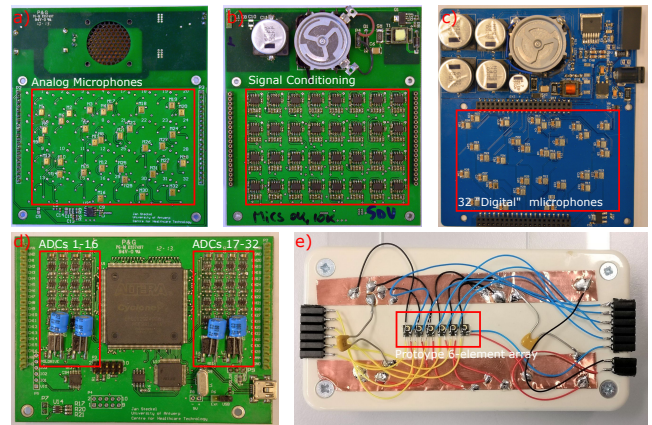


Fig. 1. Different incarnations of the proposed acoustic imaging system. Panels a, b and d show the system using analog microphones from [4], [5], containing 32 MEMS microphones with analog output, 32 signal conditioning blocks and 32 ADCs. Panel c) shows the an array with the same functionality, consisting only of 32 1-bit MEMS microphones, demonstrating a significant reduction of components (20 to 1 component per channel). Finally, panel e) shows the linear array of six Knowles SPH0641 microphones which is used in the experiments in this paper. The average inter-element distance here is around 3.4 mm, this allows for a maximum frequency of around 50 kHz.

## II. ACOUSTIC IMAGING

We wish to apply the developed microphone arrays in an 3D acoustic imaging application, focussing on the construction

of so-called energyscape representations of the environment. The energyscape is a voxel-based image-like representation of the received energy for each of the range-angle voxels. A single omnidirectional emitter emits a broadband acoustic signal (typically an frequency-modulated sweep from 80kHz to 20kHz), which is reflected by the environment and recorded using the microphone array. Afterwards, through a matched-filter process, followed by beamforming and subsequent envelope detection, an acoustic image of the environment can be created. A more detailed explanation of the signal processing behind the construction of these energyscapes is given in [4], [5], and a schematic overview of this process is shown in figure 3.

To demonstrate the applicability of the proposed system architecture for acoustic imaging, we constructed a small microphone array consisting of six Knowles SPH0641LM4H-1 microphones. Using this small array it is possible to determine if the microphones are suitable for our final application, a 32-microphone array capable of 3D imaging [1], [6]. The signal used for the proposed application is a broadband frequency modulated sweep that goes from 80 kHz to 20 kHz in three milliseconds, shown in figure 2.

### III. HARDWARE PLATFORM

The hardware architecture for efficient interfacing of these microphones is built around an FPGA-SoC, which combines an FPGA with one or more powerful ARM cores. This architecture enables capturing the one-bit data at high speeds in a synchronous fashion using the GPIO pins connected to the FPGA portion of the SoC. We use the DE0-Nano-SoC, based around a Cyclone V SoC FPGA, which combines a flexible FPGA with a dual core ARM Cortex A9, running at clock speeds of 925 MHz maximum frequency [7]. Using the AXI bridges on the chip, this data can then be written to the DDR RAM where the dual core ARM processor can access it to perform additional signal processing steps. To further improve the system, the most time-consuming DSP calculations can take advantage of hardware acceleration using the FPGA. For example, to transform the 1-bit PDM microphone data (sampled at 4.5 MHz) into 16 bit PCM data (sampled at 450kHz) an IIR low-pass filter and subsequent decimation needs to be evaluated for each microphone channel. This process can be very time-consuming on a processor architecture, as each channel needs to be processed in a serial fashion. Implementing these demodulation filters on the FPGA fabric allows processing all the channels in parallel, greatly improving the efficiency of the data-acquisition system. Because of the tightly connected Hardware Processing System (HPS) and FPGA a fast and flexible system can be created. The architecture of the hardware platform used in this paper is shown in figure 3. An advantage of these SoC devices for applications like these is that the system can almost serve as "black box" system from the eye of the user. This way the user can operate a system that delivers the data needed for the task at hand, without the need for further external processing using another computer. This leads a compact system that greatly

improves the user experience and will almost resemble a plug-and-play solution for complex tasks like sonar navigation or acoustic imaging.

As stated before, the signals originating from the microphones, connected to the FPGA, are a one-bit PDM signals, which need to be converted into 16bit PCM format, which can be achieved using an (IIR) low-pass filter. In case of the SPH0641LM4H-1 microphones, the featured sigma-delta modulator (sampled at 4.5MHz) performs noise-shaping to place the discretization noise into the higher spectral region (starting from approximately 150kHz), as can be seen in figure 2. After the use of the low pass filter, the noise on this signal is removed (figure 2, so that an accurate representation of the original signal is obtained. The output of the lowpass filter gets decimated by a factor 10, to achieve an overall sampling rate of 450kHz.

### IV. RESULTS

The experimental set-up used to get these results is similar to the set-up used in [8]. We mounted the small prototype array on a FLIR TPU46 pan/tilt system and a Senscomp 7000 transducer is used to emit the frequency modulated sweep. The data-acquisition is performed using the DE0-nano-SoC using the system architecture presented in figure 3. The data-processing (ie. energyscape generation) was performed in Matlab, but it should be noted that this processing can easily be performed by the on-board ARM cores, as shown in the architecture of figure 3. We placed the emitter in four different positions: -30, -10, 10 and 30 degrees with respect to the microphone array at a distance of 1.7 meter, and we recorded an emission of the source. Next, we calculated the energyscape for that measurement, and the resulting images are shown in figure 2. From these energyscapes, it becomes clear that the microphones and overall system concept are well suited for an acoustic imaging application. The detected sound source is clearly visible on the plots, along with a few reflections from the environment.

### V. CONCLUSION

In this paper, we presented a system concept for creating low-cost MEMS microphone arrays for ultrasonic imaging in air, consisting of microphones with a 1-bit digital interface connected to an FPGA-SOC device. The combination of the FPGA-SOC with the digital microphones can greatly reduce the overall component count of the imaging system, while retaining the same performance as a system we have previously developed using separate analog amplifiers and ADC-stages. Indeed, in relation with the analog counterpart, a decrease from 20 components per channel (microphone + conditioning + ADC-stage) to one component (microphone) can be obtained. This makes the design of a system using the microphones more attractive and cost-friendly, and better allows scaling to larger microphone arrays. The digital nature of the SPH0641 interface is optimally suited to use FPGA-based hardware such as the Altera Cyclone V. These FPGA-based processing platforms are ideally suited to perform the synchronous acquisition and

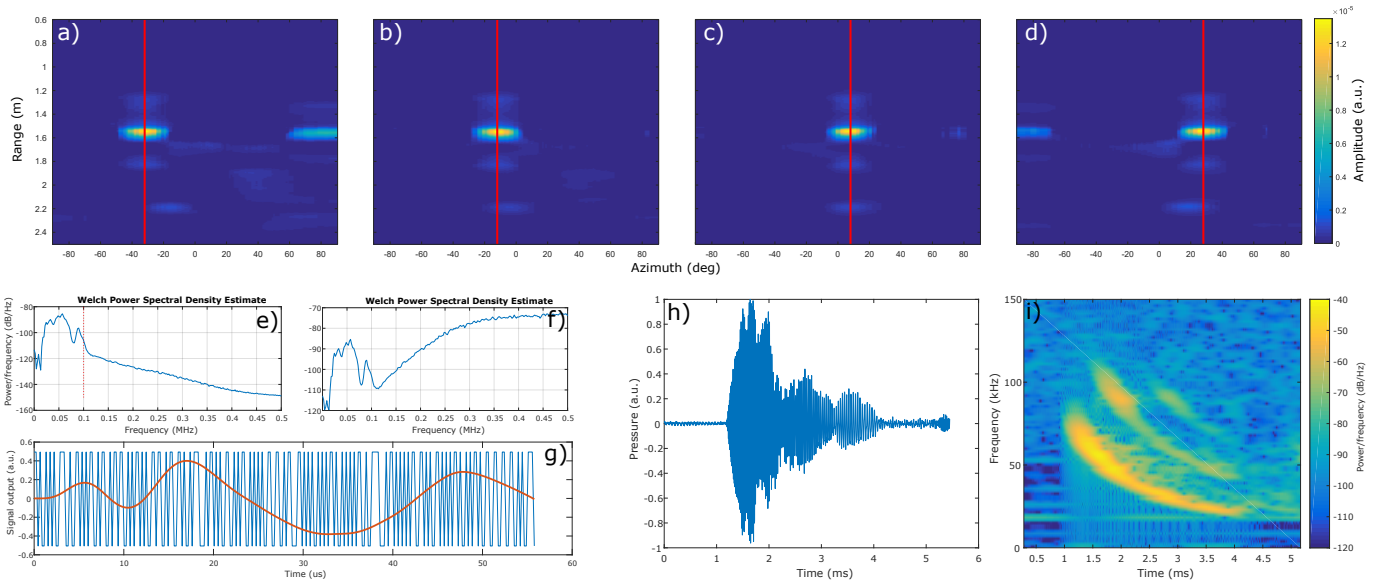


Fig. 2. Overview of the results obtained in this paper. Panels a-d show the energyscapes obtained with the linear array with an acoustic source placed at -30, -10, 10 and 30 degrees. The red line indicates the true angle for the source. These energyscapes indicate clearly the position of the source, and several reflections from the environment can also be resolved, indicating the applicability of the proposed microphones. Panels e) and f) show the noise shaping circuitry of the PDM microphones and the effect of the lowpass filtering on the acoustic spectrum. The spectra are truncated at 500kHz, but extend to 2.25MHz (half the PDM sample rate). Panel g) shows the raw 1-bit PDM signal (blue line) and the signal after IIR low-pass filtering (red line). Panel h) shows the recorded time-pressure signal for one microphone, with panel i) showing the spectrogram of this time-pressure signal. This spectrogram shows the effective acoustic range of the used microphone, well above 100kHz.

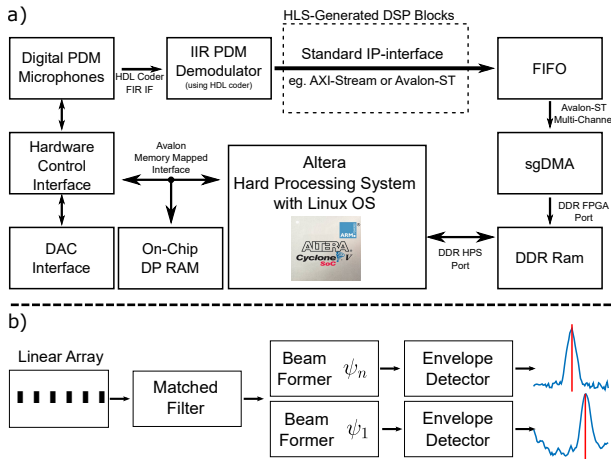


Fig. 3. Schematic overview of the system architecture. Panel a) overview of the proposed hardware platform, consisting of a DE0-Nano-SoC board, featuring a Cyclone V SoC. The FPGA-soc enables fast connectivity between the FPGA fabric and the on-board ARM Dual Core processor and the DDR memory. With this architecture data can rapidly be stored coming from the GPIO pins on the FPGA side of the SoC, while the Hard Processing System (HPS) handles user communication and data processing. Time consuming calculations (like IIR filtering at high sampling rates and decimation) can also take advantage of hardware acceleration on the FPGA. Panel b) shows the different steps used in the signal processing algorithm to create the energyscapes. After matched-filtering, the beamformers form an acoustic lens, making the array sensitive for a chosen direction. After beamforming, envelope detection extracts the energy-profile in function of range, and several of these range-energy profiles are combined to form an acoustic image, called the energyscape.

demodulation of many microphone signals, which would be very time-consuming on traditional processor-based hardware platforms. The resulting imaging system performs equally well as the analog counterpart, demonstrating the applicability of these novel microphone architectures.

## REFERENCES

- [1] J. Steckel, A. Boen, and H. Peremans, "Broadband 3-d sonar system using a sparse array for indoor navigation," *IEEE Transactions on Robotics*, vol. 29, no. 1, pp. 161–171, 2013. [Online]. Available: <http://ieeexplore.ieee.org/document/6331017/>
- [2] Knowles Acoustics, "Knowles Acoustics", Std. [Online]. Available: <http://www.knowles.com>
- [3] —, "SPH0641 Datasheet", Std. [Online]. Available: <http://www.knowles.com>
- [4] J. Steckel and H. Peremans, "Spatial sampling strategy for a 3d sonar sensor supporting batslam," *Intelligent Robots and Systems (IROS)*, 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/7353452/>
- [5] —, "Sparse decomposition of in-air sonar images for object localization," *IEEE Sensors*, 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6985263/>
- [6] —, "Acoustic flow-based control of a mobile platform using a 3d sonar sensor," *IEEE Sensors Journal*, vol. 17, pp. 3131–3141, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/7888490/>
- [7] Altera, Intel Corporation, "Cyclone V SoC datasheet", Std. [Online]. Available: [https://www.altera.com/en\\_US/pdfs/literature/hb/cyclone-v/cv\\_51002.pdf](https://www.altera.com/en_US/pdfs/literature/hb/cyclone-v/cv_51002.pdf)
- [8] R. Kerstens, D. Laurijssen, J. Steckel, and W. Daems, "Widening the directivity patterns of ultrasound transducers using 3-d-printed baffles," *IEEE Sensors Journal*, vol. 17, pp. 1454–1462, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/7792174/>