

HARDWARE AND SOFTWARE FOR REPRODUCIBLE RESEARCH IN AUDIO ARRAY SIGNAL PROCESSING

Eric Bezzam*, Robin Scheibler*, Juan Azcarreta*, Hanjie Pan*, Matthieu Simeoni*[†],
René Beuchat*[‡], Paul Hurley[†], Basile Bruneau*, Corentin Ferry*, and Sepand Kashani*[†]

*École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

[†]IBM Zurich Research Laboratory, Switzerland

[‡]Haute École du Paysage, d'Ingénierie et d'Architecture de Genève, Switzerland

ABSTRACT

In our demo, we present two hardware platforms for prototyping audio array signal processing. *Pyramic* is a 48-channel microphone array fitted on an FPGA and *Compact Six* is a portable microphone array with six microphones, closer to the technical constraints of consumer electronics. A browser based interface was developed that allows the user to interact with the audio stream from the arrays in real time. The software component of this demo is a Python module with implementations of basic audio signal processing blocks and popular techniques like STFT, beamforming, and DoA. Both the hardware design files and the software are open source and freely shared. As part of a collaboration with IBM Research, their beamforming and imaging technologies will also be portrayed.

The hardware will be demonstrated through an installation processing the microphone signals into light patterns on a circular LED array. The demo will be interactive and let visitors play with different algorithms for DoA (SRP, FRIDA [1], Bluebild) and beamforming (MVDR, Flexibeam [2]).

The availability of an open platform with reference implementations encourages reproducible research and minimizes setup-time when testing and benchmarking new audio array signal processing algorithms. It can also serve as a useful educational tool, providing a means to work with real-life signals.

Index Terms—Microphone array, audio signal processing, beamforming, DoA, spectral imaging.

1. INTRODUCTION

Microphone arrays are increasingly prominent in popular consumer electronics, such as the Amazon Echo (7 microphones), Google Home (2 microphones), Apple AirPods (2 microphones), and high-end smartphones (2 microphones). These devices rely on *beamforming*, i.e. coherently adding signals from more than one microphone to improve the signal-to-noise ratio (SNR) of a desired signal. Furthermore, it is customary to precede beamforming by *direction-of-arrival* (DoA) estimation in order to spatially locate the desired signal. Microphone arrays provide more robust indoor localization and tracking [3], which is especially beneficial for acoustic SLAM in robotic applications, dereverberation in speech detection [4] and room geometry reconstruction [5]. The applications above would all greatly benefit from a

convenient and portable platform as proposed in this work. When testing new algorithms, it allows one to circumvent the cumbersome setup (and takedown) typically required: a bulky audio interface connected to a computer, an XLR cable from each microphone to the audio interface, and perhaps multiple microphone stands. Moreover, interfacing the microphone array with a portable processing unit such as a field programmable gate array (FPGA) or a single-board computer allows for a certain amount of processing to be done on the device itself. Such setups make single cable and/or wireless communication feasible with a host computer. Removing the complications of setup and providing a convenient interface for acquisition and processing provides an accessible means for testing under a variety of conditions. Moreover, the availability¹ of open source hardware and software modules facilitates reproducible research and fosters collaborations between academia and industry, as successfully illustrated by our current association with IBM.

A microphone array with a convenient interface can also serve as an educational tool, providing users a hands-on experience with signal processing on measurements perceptible to human senses – audio. A graduate course entitled *Audio signal processing and virtual acoustics* is already offered by the Audio-visual Communications Laboratory (LCAV) at EPFL, and such a platform could complement the theory with practical insight.

At LCAV, we have developed two platforms for the aforementioned purposes: the *Pyramic* (in cooperation with the Processor Architecture Laboratory at EPFL) and the *Compact Six*, which can both be seen in Fig. 1. The ultimate goal is to create an open platform in order to promote reproducible research and the development of educational tools in computational acoustics.

2. HARDWARE PLATFORMS

The main challenges when designing microphone arrays arise when considering the trade-offs between array weight, size, power consumption, cost and flexibility. Although using more microphones typically improves performance, it compromises the flexibility of the system. In order to tackle these constraints, the *Pyramic* and *Compact Six* arrays take advantage of recent advances in MEMS technologies, providing us with small high-quality microphones at an affordable price.

IBM acknowledges the funding by the Dutch Ministry of Economische Zaken, and the province of Drenche as part of the DOME project.

¹The software and hardware designs can be found at <http://github.com/LCAV>

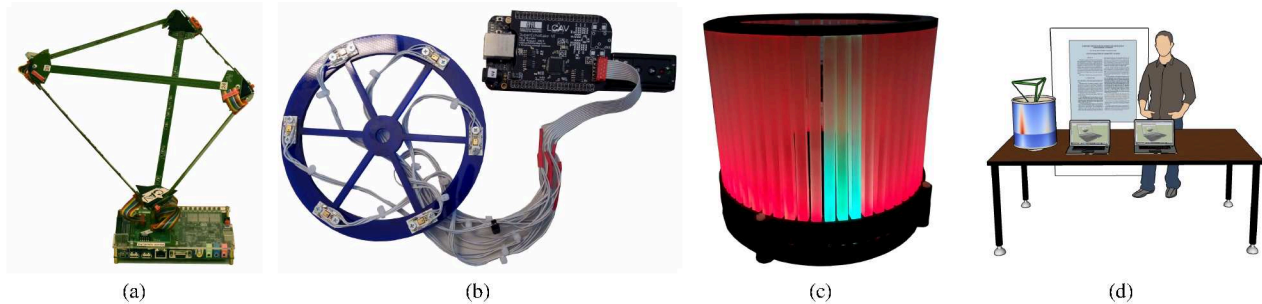


Fig. 1: (a) Pyramic and (b) Compact Six microphone arrays. (c) Circular LED monitor. (d) ICASSP demo setup.

2.1. Pyramic 48-channel Microphone Array

The Pyramic array is assembled from six linear arrays into a triangular pyramid, hence the name *Pyramic* [6]. Each linear array is 30 cm long and contains eight microphones of varying spatial resolution. The former are connected to an Altera FPGA and can communicate with any SSH-capable host computer via Ethernet [7]. Note that the six linear arrays can be reassembled into a different shape if desired.

2.2. Compact Six Microphone Array

The Compact Six array comprises of six microphones uniformly spaced along a circular disk. The shape of the array can also be changed by laser cutting a new printed circuit board. The microphone array interfaces with a BeagleBone Black (BBB) [8] running Linux (Debian Wheezy) and can communicate with a host computer via Ethernet or USB. It is significantly cheaper (around \$55) and more portable than the Pyramic.

3. SOFTWARE MODULE

In addition to the hardware platforms, we have a comprehensive Python package that can readily be used on recordings. The package contains standard processing tools, popular beamforming and DoA algorithms, as well as some more recent algorithms from our own research [1, 9, 2].

4. PROPOSED DEMO

During the Demo session at ICASSP 2017, we would like to demonstrate the DoA, imaging, and beamforming capabilities of the Pyramic and Compact Six arrays. Specifically, the performance of the FRIDA algorithm (described in the ICASSP 2017 submission [1]) will be compared with the standard DoA methods SRP-PHAT, MUSIC, and TOPS. Moreover, the IBM Bluebird imaging algorithm and its multi-resolution capabilities will be shown in action. On the beamforming side, various algorithms will be investigated, including the recent Flexibeam algorithm which allows one to achieve a wide range of analytically-specified radiation patterns [2].

For visualization purposes, the Pyramic array will be equipped with a circular LED monitor, which can display the DoA of multiple sources. This is done by mapping the response spectrum

of a particular DoA algorithm to a specific set of colors, e.g. green for higher values, which correspond to estimated source directions, and red for lower values. As part of the demo, users will also be able to play with the aforementioned technologies through the interactive browser interface.

Thanks to the portability of the Pyramic and Compact Six microphone arrays, we would only require a table for two laptops and the Pyramic (30 cm-sided tetrahedron) for our demonstration. A rough sketch of our planned setup can be seen in Fig. 1.

5. REFERENCES

- [1] H. Pan, R. Scheibler, E. Bezzam, I. Dokmanić, and M. Vetterli, “FRIDA: FRI-based DoA estimation for arbitrary array layouts,” 2016, submitted to ICASSP 2017.
- [2] P. Hurley and M. Simeoni, “Flexibeam: Analytic spatial filtering by beamforming,” in *Proc. IEEE ICASSP*, Shanghai, 2016, pp. 2877–2880.
- [3] I. Marković and I. Petrović, “Speaker localization and tracking with a microphone array on a mobile robot using von mises distribution and particle filtering,” *Robot. Auton. Syst.*, vol. 58, no. 11, pp. 1185–1196, Nov. 2010.
- [4] P. Naylor and N. Gaubitch, *Speech Dereverberation*, ser. Signals and Communication Technology. Springer London, 2010.
- [5] I. Dokmani, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, “Acoustic echoes reveal room shape,” *Proc. Natl. Acad. Sci. USA*, vol. 110, no. 30, pp. 12 186–12 191, 2013.
- [6] J. Azcarreta Ortiz, “Pyramic array: An FPGA based platform for many-channel audio acquisition,” Master’s thesis, EPFL, Lausanne, Switzerland, Aug. 2016.
- [7] “DE1-SoC Board,” <http://www.terasic.com.tw/cgi-bin/page/archive.pl?Language=English&No=836>, accessed: 2016-10-14.
- [8] “Beaglebone Black,” <http://beagleboard.org/black>, accessed: 2016-10-13.
- [9] I. Dokmanić, R. Scheibler, and M. Vetterli, “Raking the cocktail party,” *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 5, pp. 825–836, 2015.