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, Switzerland †IBM Research — Zürich ‡Haute École du Paysage, d'Ingénierie et d'Architecture de Genève, Switzerland*

Benefits of Open HW/SW Framework

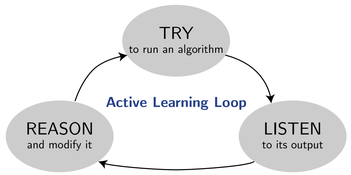
There are significant challenges to go from theory to experiments. We present a set of software and hardware tools that streamline this process.

Reproducible Research

- Reduce *time-to-market* for new research
- Saves implementation time
- Provides uniform benchmark
- Simple, affordable, and repeatable experimental setup

Education

- Easier setup saves time in class
- Gain intuition about the nature of sound and processing
- Go from theory to practice in minutes

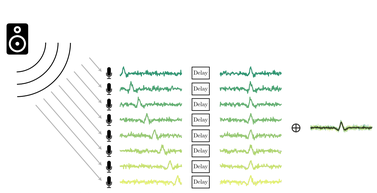


Pyroomacoustics

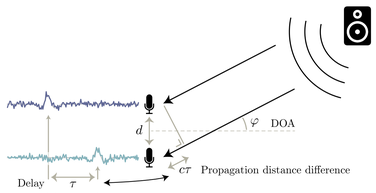
A **Python** package for array signal processing

- Multiplatform, free, and open source
- High performance linear algebra via `numpy/scipy`

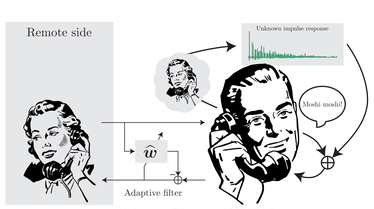
Features



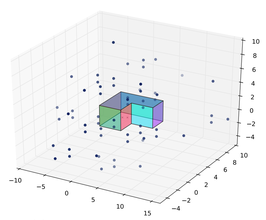
(a) Beamforming (Delay-and-Sum, MVDR)



(b) DOA (MUSIC, SRP-PHAT, CSSM, TOPS, WAVES)

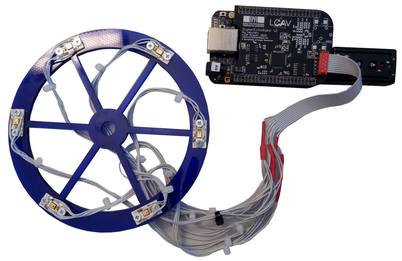


(c) Adaptive Filtering (LMS, RLS)



(d) High performance RIR generator

Hardware I: CompactSix



- 6 MEMS microphones, free geometry
- Sampling frequency 48 kHz or 192 kHz
- Beaglebone Black platform
- ARM CPU @ 1GHz CPU, 512 MB RAM
- ALSA drivers

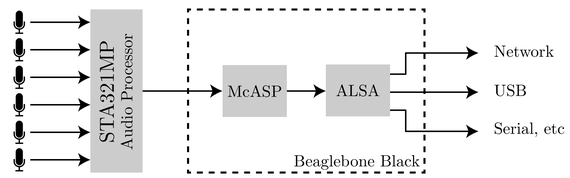
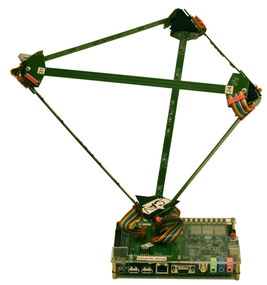


Figure 1: System architecture of CompactSix

Hardware II: Pyramic

- 48 MEMS microphones on 6 PCBs
- 2 output channels
- Sampling frequency 48 kHz
- System-On-Chip with FPGA/ARM
- FPGA reads samples into shared memory
- ARM CPU @ 800 MHz, 1 GB DDR3 RAM
- Shared library `libpyramicio.so`



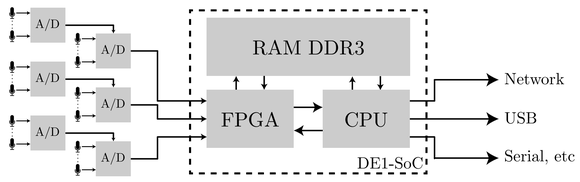
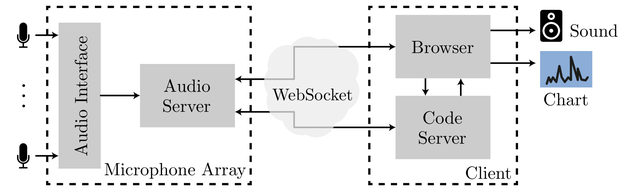


Figure 2: System architecture of Pyramic

EasyDSP: Browser-based Interface

- Intuitive interface
- Interactive processing of the audio stream
- Connection through network
- Multiplatform, based on web technologies (JavaScript, WebSocket)
- Processing in Python

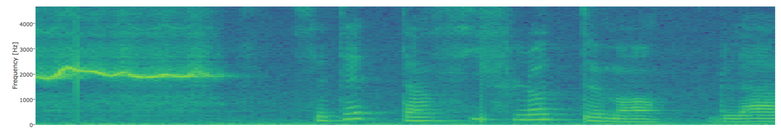
Modular architecture



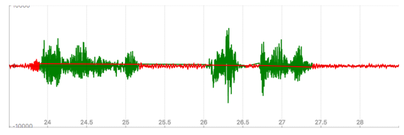
Easily extendable to other Linux-based microphone arrays

Live Plots and Audio

Data can be sent back to the browser for live plotting.



(a) Spectrogram



(b) Speech detection

(c) Direction of arrival

Publications using the Framework

H. Pan et al., *FRIDA: FRI-based DOA Estimation with Arbitrary Array Layout*, 2017
(Presented last Tuesday!)
R. Scheibler, M. Vetterli, *The Recursive Hessian Sketch for Adaptive Filtering*, 2016
I. Dokmanić, R. Scheibler, M. Vetterli, *Raking the Cocktail Party*, 2015
R. Scheibler, I. Dokmanić, M. Vetterli, *Raking Echoes in the Time Domain*, 2015

Get it all!

Free, as in *free speech*, and as in *free beer*!!
Everything is available at `http://github.com/LCAV/`
Repos: CompactSix, Pyramic, pyroomacoustics, easy-dsp

We would like to acknowledge the kind help of André Guignard with the fabrication of the LED-based demonstration device. We also thank Sahand Kashani for his contributions to the design of the Pyramic array and making it compatible with the browser interface.