



**SKFM 2021** 

Studentská konference Fakulty mechatroniky, informatiky a mezioborových studií

# **Semi-Blind Speaker Extraction Performed On-line** on Dense Microphone Array

J. Čmejla, T. Kounovský, J. Janský, J. Málek, M. Rozkovec and Z. Koldovský <jaroslav.cmejla@tul.cz> | <tomas.kounovsky@tul.cz> | <jakub.jansky@tul.cz>

We present an experimental device for speaker extraction and physical tracking and demonstrate its use in real conditions. The device is equipped with a dense planar array consisting of 64 microphones mounted on a rotating platform. Blind source extraction algorithms controlled by x-vector piloting are used to extract the desired speaker, which can be physically tracked by the rotating microphone array. Keywords: source extraction, source tracking, x-vectors, Independent Vector Extraction

### **Motivation**

- Speaker extraction and tracking in difficult conditions
  - Multiple speaker cross-talk
  - Background noise
- Moving sources

## Device

- 64 MEMS microphones in a grid pattern
- $(2 \times 2 \text{ cm equidistant spacing})$
- FPGA controller for data acquisition and transfer



Experimental device with a huge number of microphones (Fig. 2)

#### Algorithm

- **Extraction algorithms** 
  - CSV-AuxIVE [1]
  - QuickIVE [2]
  - FastDIVA [3]
- **CSV model** to allow for source movement [4]
- Piloting for assured convergence
  - X-vector TDNN for feature extraction [5]
  - PLDA classifier to compare new features with known speaker features (enrollment set)



- **3D printed** rotary platform
- Unlimited rotation
- NEMA 17 stepper motor, TMC2208 silent stepper driver and timing belt for silent operation



Figure 2: 64 microphone array board with a rotary platform. 20×20 cm version of the board.

#### **Software**

- Written in MATLAB
- Handles communication with device, data acquisition, source extraction and tracking

Track
Start
1. mic

• Graphical user interface (Fig. 1) allows for easy control of the system

Allows adding new speakers to the enrollment set on-the-fly

#### **On-line DEMO**

**Figure 1:** Device control software interface (Matlab GUI).

**On-line audio & video stream** from A203 (formerly known as SMART Room) 

#### References Acknowledgment

This work was supported by the Student Grant Scheme of the **Technical University of Liberec** (2021).

- [1] J. Janský et al. Auxiliary function-based algorithm for blind extraction of a moving speaker. arXiv 2002.12619, 2020.
- Z. Koldovský et al. Algorithm for independent vector extraction based on semi-time-variant mixing model, arXiv 1910.10242, 2021.
- Z. Koldovský et al. Dynamic independent component/vector analysis: Time-variant linear mixtures separable by time-invariant beamformers, IEEETransactions on Signal Processing, 2021.

N. Amor et al. Blind extraction of moving sources via independent component and vector analysis: Examples, in *ICASSP 2021*.

J. Janský et al. Adaptive blind audio source extraction supervised by dominant speaker identification using x-vectors, in ICASSP 2020. [5]