Spike-based Beamforming using pMUT Arrays for Ultra-Low Power Gesture Recognition

Emmanuel Hardy
System Overview

› End-to-end Ultrasonic Gesture Recognition with pMUTs

› Competing technologies: IR, electric field, radar, camera.

› Key advantages:
  ■ Insensitivity to light/shadow
  ■ Low cost
  ■ Low power
Key Points

› End-to-end Power Efficient Edge AI device
  ■ Efficient sensors
  ■ Analog-to-information conversion
  ■ Low power inference

› Needs a widely cross-functional team
The team

› Bruno Fain
› Thomas Mesquida
› François Blard
› François Gardien
› François Rummens
› Jean-Claude Bastien
› Jean-Rémi Chatroux
› Sébastien Martin
› Venceslass Rat
› Elisa Vianello

doi: [10.1109/IEDM45625.2022.10019395](https://doi.org/10.1109/IEDM45625.2022.10019395).
1. Sensor Array
PMUTs Transducers

PMUT: Piezoelectric Micromachined Ultrasonic Transducer

› Aluminum nitride piezo material
› Bimorph structure for higher sensitivity and TX power
› Resonance ~ 100kHz (tunable)
› 8” MEMS production line in CEA-Leti
Acoustic setup

- **TX**: 1 PMUT
- **RX**: 2 orthogonal arrays of 5 PMUTs
- **3D sensing**
- **Range**: 10 -> 60 cm
2. Spike-based signal processing
Spike-based Beamforming

\[ \delta_t = \frac{d \sin \alpha}{c} \]

Direction of arrival

Time Difference of Arrival between M_0 and M_1

Conversion to spikes
Spike-based Beamforming

\[ \delta_t = \frac{d \sin \alpha}{c} \]

Direction of arrival

Time Difference of Arrival between M_0 and M_1

Spike Coherence Detection

\[ \sum S_n \]

Direction 0°

Threshold

Window length

Direction \( \alpha \)

\[ d \sin \alpha \]

\[ M_0 \quad M_1 \quad M_2 \]

\[ d \]

\[ c \]

\[ 2\delta_t \]

\[ \delta_t \]
Spike-based Beamforming

› 11 directions
› +/-50° range with 10° steps
› Optimal threshold and window size
Spike-based Beamforming

- wrt. Conventional Beamforming
  - ~3 times better angular selectivity at 0°
  - No side lobes
  - Increased range

- BUT not proportional to signal amplitude.
Feature Vector

Coherence Matrix y

Coherence Matrix x

Distance Vector

Direction Vector x

Direction Vector y

Feature $x_t$ 35x1
Example of Gestures

[Diagrams and graphs showing angle x, angle y, and distance over frames]
3. Classification & Results
**Gesture Classifier**

› **Recurrent Neural Network**
  - Trainable Temporal Dependency (no fixed window)
  - Computational Efficiency

› **Two classifiers:**
  1. GRU Baseline
  2. Spiking RNN Hardware Target
Spiking Recurrent Unit

- $V(t) = x(t) - V_{th}$
- $y(t) = \begin{cases} 1 & \text{if } V(t) > V_{th} \\ 0 & \text{otherwise} \end{cases}$

Input spikes

Membrane potential

Output spikes

$V = x(t) - V_{th}$

$t_{FRAME}$

Threshold

Leak

Delay

Refractory period
Right-Left example

4 bits Feature Quantization
Experimental Setup

- **Beam-forming (FPGA)**
  - pMUT RX
  - pMUT TX
  - Discrete Front-end Electronics
  - Converters
  - Driver
  - 12 bits
  - 500kHz
  - SPI
  - Spikes
  - Classes
  - SRNN
Gesture Dataset

› 5 gestures + None class

› 499 examples, 12 participants
  ■ Training/test (9/3 split)

› 10 to 50 cm distance

› Data augmentation for training
  ■ System symmetry
Classification Results

Small accuracy drop with SRNN

<table>
<thead>
<tr>
<th>Action</th>
<th>GRU-16</th>
<th>SRNN-110</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeftRight</td>
<td>87.2%</td>
<td></td>
</tr>
<tr>
<td>RightLeft</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upwards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Downwards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PushPull</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>86.0%</td>
<td></td>
</tr>
</tbody>
</table>
### Classification Results

#### Gesture ambiguity, small dataset

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LeftRight</td>
<td>RightLeft</td>
</tr>
<tr>
<td>LeftRight</td>
<td>72.4%</td>
<td>0.0%</td>
</tr>
<tr>
<td>RightLeft</td>
<td>0.0%</td>
<td>93.8%</td>
</tr>
<tr>
<td>Upwards</td>
<td>0.0%</td>
<td>4.2%</td>
</tr>
<tr>
<td>Downwards</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>PushPull</td>
<td>1.6%</td>
<td>0.0%</td>
</tr>
<tr>
<td>None</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

tinyML EMEA Innovation Forum | 28th June 2023
## State-of-the-Art

<table>
<thead>
<tr>
<th></th>
<th>This work</th>
<th>Przybyla et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of transducers</strong></td>
<td>TX-RX: AlN pMUT</td>
<td>TX-RX: AlN</td>
</tr>
<tr>
<td><strong># RX - Pattern</strong></td>
<td>10 - 2 Lines X/Y</td>
<td>7 - Zigzag</td>
</tr>
<tr>
<td><strong>Classif. type</strong></td>
<td>SRNN</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>86.0% (5 gest.)</td>
<td>64.5% (5 gest.)</td>
</tr>
<tr>
<td><strong>Meas. period</strong></td>
<td>40 ms</td>
<td>5.9 ms</td>
</tr>
<tr>
<td><strong>Max. range</strong></td>
<td>60 cm</td>
<td>100 cm</td>
</tr>
<tr>
<td><strong>Hardware integration</strong></td>
<td>COTS</td>
<td>Post-process</td>
</tr>
<tr>
<td><strong>Est. sensing energy (ASIC)</strong></td>
<td>78.1 nJ/meas.</td>
<td>15.6 µJ/meas.</td>
</tr>
<tr>
<td><strong>Est. inference energy (ASIC)</strong></td>
<td>330/760 nJ/meas. (None/Gesture)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Wrapping Up...
Applications and Perspectives

› Gesture Recognition
   ■ Wearables, automotive, VR headsets.

› Robotics
   ■ Obstacle detection
   ■ Beamforming at emission

› What’s Next
   ■ Apply this approach to new sensors from CEA Leti
   ■ Build new exciting prototypes and ASICs
Takeaway Points

1. Small form factor pMUT array
2. Beamforming & Signal Processing in the Spike domain
3. Low Power Gesture Classification

Analog-to-Information strategy to yield more efficient Sensors + Edge AI systems.
Thank you!

Emmanuel Hardy
emmanuel.hardy@cea.fr
CEA Leti
17 avenue des martyrs
38000 Grenoble
FRANCE
Copyright Notice

This presentation in this publication was presented as a tinyML® EMEA Innovation Forum. The content reflects the opinion of the author(s) and their respective companies. The inclusion of presentations in this publication does not constitute an endorsement by tinyML Foundation or the sponsors.

There is no copyright protection claimed by this publication. However, each presentation is the work of the authors and their respective companies and may contain copyrighted material. As such, it is strongly encouraged that any use reflect proper acknowledgement to the appropriate source. Any questions regarding the use of any materials presented should be directed to the author(s) or their companies.

tinyML is a registered trademark of the tinyML Foundation.

www.tinyml.org