DOI 10.5162/EUROSENSORSXXXVI/PT2.210

Towards fully hardware-based neuromorphic encoding for efficient vibration signal recognition

Tamás Zeffer[1,2], Tímea Nóra Török[3,1], László Pósa[1], Ferenc Braun[1], András Halbritter[3], János Volk[1]
[1] HUN-REN Centre for Energy Research, 1121 Budapest, Hungary
[2] Doctoral School on Material Sciences & Technologies, Óbuda University, 1034 Budapest, Hungary
[3] Budapest University of Technology and Economics, 1111 Budapest, Hungary

Corresponding Author's e-mail address: volk.janos@ek.hun-ren.hu

Summary:

Neuromorphic signal processing can enhance the efficiency of IoT sensors and support edge computing solutions. However, the preprocessing of the encoded signals, to be transferred to a spiking neural network, requires high computational power. In this work, we propose an energy-efficient hardware-based solution for the analysis of rapidly changing vibration or acoustic signals. This was inspired by the human cochlear implant, which exploits the plasticity of the human brain to enable clear speech recognition even on a very limited number of frequency channels (16-22). Our proposed hardware consists of a 16-channel frequency-selective MEMS cantilever array, and a VO_2 memristor nanogap based oscillator for amplitude sensitive spiking signal generation. To test our solution, we used Google Command Speech benchmark database.

Keywords: Spiking neural network, cantilever array, memristor oscillator, vibration analysis, edge computing

Motivation and Objective

For quite some time, it has been evident that biological systems surpass electronic counterparts in terms of energy efficiency, highlighting a significant gap. Therefore, the development of neuromorphic auditory hardware systems holds promise for achieving greater efficiency and performance in future hardware audio solution [1].

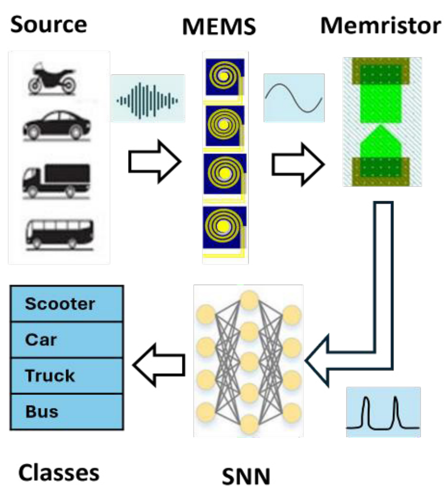

Fig. 1. Proposed SNN-based vibration source classification method using piezo-MEMS frequency sensitive cantilever array and memristor to encode amplified analogue signals into neural spikes.

In this paper, we propose a solution for FFT free neuromorphic encoding of acoustic and vibration spectrograms. It is based on our previous work, where similar cantilever arrays were used for a fully implantable cochlear implant [2]. Fig. 1 shows an application example of an Acoustic Vehicle Detection system. Specifically, MEMS cantilevers serve as sensors to detect the sound or seismic vibrations emitted by different vehicles. These cantilevers are tuned to specific resonant frequencies. Subsequently, the harmonic motion of each piezoelectric cantilever generates voltage outputs, which are processed by oscillatory circuitry to produce spikes. These circuits incorporate VO_2-based memristors, which play a crucial role in spike generation. The resulting spikes serve as inputs to a Spiking Neural Network (SNN), enabling the classification.

Description of the new method

For a concise SNN network, we have chosen the speech2spikes pipeline [3], which consists of blocks that initially utilize the Mel Spectrogram, followed by their Step-Forward and eventually by Cumulative Sums, to produce the input spikes to the SNN.

To achieve higher accuracy in hardware system design, we modeled our existing hardware elements in two distinct steps. Our models are based on the fabricated cantilever and memristor along with its experimental signals (Fig. 2) Initially, we aligned the data of our 4x4 spiral cantilevers to the parameters of the mel-spectrogram, changing the number of bins from 20 to 16, and reducing the frequency range to a minimum and a maximum frequency of 200 and 700 Hz [4].

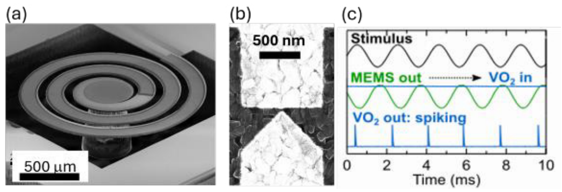

Fig. 2. Fabricated spiral shaped piezocantiver having tuned resonant frequency (a), the VO2 nanogap memristor (b), and the measured signal conversion (c).

Subsequently, we modeled the memristive oscillatory circuit, omitting the use of step-forward and cumulative sum pipeline algorithms. We summarized this two-step hardware design approach in the rows of the Table 1 below, marked as version 1 and 2 for clarity. Emulation of hardware is emphasized by the light blue background of the cells. The number of input neurons in the SNN network, determined by the bin numbers of the mel-spectrum or the count of spiral cantilevers designed for various natural frequencies. Each layer of the SNN network consists of 256 LIF neurons. Efficiency is measured in terms of accuracy.

Version	Bins	Frequencies	Methods
0	20	20…20.000 Hz	step-forward and cumulated sum algorithm
1	16	200…700 Hz	
2			emulated oscillatory circuit on measurements

Table 1. Co-deisgn of preprocessing MEMS and SNN with emulated hardware models. Cells in light blue denote the emulated components.

Results

We succeeded in achieving a 40% accuracy with the first version, which may signify a good result in our work. Although this represents promising progress, further optimization and refinement are needed to advance accuracy. Training and modeling of version 2 are still in progress. Unfortunately, the accuracy of the zeroth version during training fell short of expectations, showing only 70%. However, if we manage to find the appropriate training conditions, we believe that the accuracy will improve in version 2, following the published 88%.

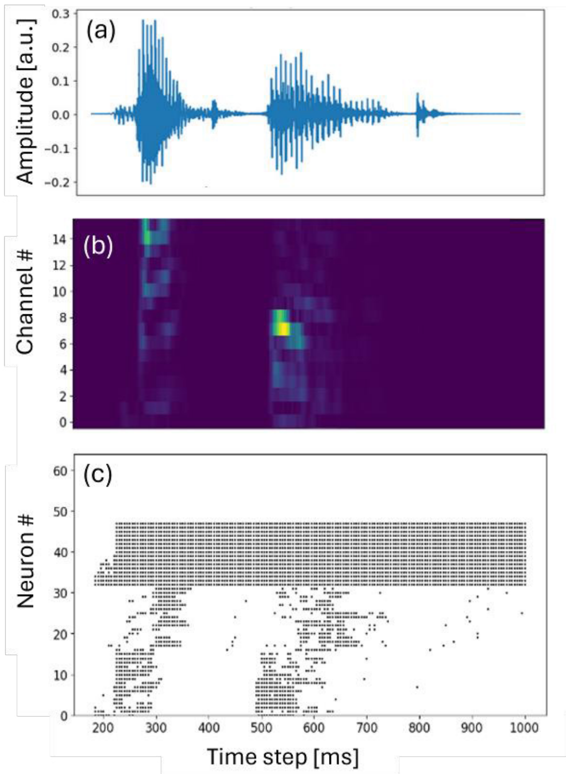

Fig. 3. (a) The preprocessing pipeline signals. Initially, the 'backward' GSC audio data (a) enters the pipeline and undergoes processing by the Mel Spectrogram algorithm with 16 bins (b). Eventually, the Step-forward and cumulative sums algorithm feeds 64 neurons of the SNN with its generated spikes (c).

References

[1] N. Kasabov *et al.*, “Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: Design methodology and selected applications,” *Neural Networks*, vol. 78, pp. 1–14, Jun. 2016, doi: 10.1016/j.neunet.2015.09.011.

[2] P. Udvardi *et al.*, “Spiral-Shaped Piezoelectric MEMS Cantilever Array for Fully Implantable Hearing Systems,” *Micromachines*, vol. 8, no. 10, p. 311, Oct. 2017, doi: 10.3390/mi8100311.

[3] K. M. Stewart, T. Shea, N. Pacik-Nelson, E. Gallo, and A. Danielescu, “Speech2Spikes: Efficient Audio Encoding Pipeline for Real-time Neuromorphic Systems,” in *Neuro-Inspired Computational Elements Conference*, San Antonio TX USA: ACM, Apr. 2023, pp. 71–78. doi: 10.1145/3584954.3584995.

[4] L. Pósa *et al.*, “Applying Neurodynamic Behavior of Mott Memristors for Auditory Sensing,” in *Proceedings of the Neuronics Conference*, València, Spain: FUNDACIO DE LA COMUNITAT VALENCIANA SCITO, Dec. 2023. doi: 10.29363/nanoge.neuronics.2024.017.