Towards the Neuromorphic Implementation of the Auditory Perception in the iCub Robotic Platform

Daniel Gutierrez-Galan
University of Sevilla
Sevilla, Spain
dgutierrez@atc.us.es

Chiara Bartolozzi
Istituto Italiano di Tecnologia
Genoa, Italy

Juan Pedro Dominguez-Morales
University of Seville
Seville, Spain

Angel Jimenez-Fernandez
University of Seville
Seville, Spain

Alejandro Linares-Barranco
University of Seville
Seville, Spain

ABSTRACT

Hearing can be considered as one of the most important senses since it plays a key role in the audiovisual learning process. While a lot of effort has been made for achieving good results from the traditional approach of auditory perception, new trends such as neuromorphic computing are showing promising achievements in the implementation of brain structures for sensory perception. In this work, the design and integration of a neuromorphic event-based digital model of the auditory ascending pathway within the iCub robotic platform is proposed. This model, which comprises from the cochlea up to the inferior colliculus, replaces the traditional approach for sound processing already implemented on iCub and is able to perform sound recognition and spatial localization in real-time, allowing the implementation of auditory attention models in complex scenarios, like the cocktail party problem.

CCS CONCEPTS

• **Computing methodologies → Bio-inspired approaches**; • **Applied computing → Event-driven architectures**.

KEYWORDS

neuromorphic auditory sensor, iCub, event-based processing, sound recognition, sound localization

1 INTRODUCTION

Humanoid robots will soon be a common presence among humans since their capabilities of interaction with the environment and humans are steadily increasing, with object and human recognition [9], and speech understanding [5]. However, these capabilities commonly result in a high computational cost and are often deployed in remote computing systems, with the need of constant data transfer. Neuromorphic perception and computation, inspired by the computational principles of biological neural systems, can be used towards low-power, embedded solutions of complex perceptive tasks. In this work, a digital event-based bio-inspired model [6] of the Auditory Ascending Pathway (AAP) [1] on the iCub humanoid platform [7] was implemented as a neuromorphic alternative to the traditional auditory models in order to reduce latency, power consumption and computational cost. This model, which has been validated for sound recognition and sound source localization, can be extended to implement more complex auditory processing and audio-visual fusion [10], in order to improve the accuracy and learning capabilities of the system.

2 GIVING ICUB THE SENSE OF HEARING

The auditory model comprises the Neuromorphic Auditory Complex (NAC), implemented on the Zynq Field Programmable Gate Array (FPGA) for reconfigurability and real-time signal processing, and the Inferior Colliculus (IC) on SpiNNaker, where larger brain areas can be implemented, as shown in Fig. 1 (A). Thanks to the bidirectional communication between iCub and SpiNNaker [8] through Yet Another Robot Platform (YARP) [2], real-time, close-loop audio applications can be implemented. The NAC module was integrated within the iCub robotic platform as an Intellectual Property (IP) core connected to the Head Processing Unit (HPU) module. The NAC receives sound input from the two microphones placed in the ears of the iCub. It is composed of three sub-modules: 1) a 32-channels binaural Neuromorphic Auditory Sensor (NAS) [6], which implements an event-based digital cochlea model; 2) the Spike-based Superior Olive (SSO), which implements both the Interaural Time Difference (ITD) [3] and the Interaural Level Difference (ILD) extraction from the NAS' output events; and 3) the events monitor together with the interface to the HPU core, which is carried out by using the Address Event Representation (AER) protocol. The sound features extracted with the NAC are combined and integrated by a multilayer Spiking Neural Network (SNN) in SpiNNaker that models the behavior of the IC [1] to extract the relevant information regarding the location of the object in the horizontal axis. Then, the output of the SNN is sent back to the iCub to turn its head towards the sound source. This close-loop system allows real-time sound

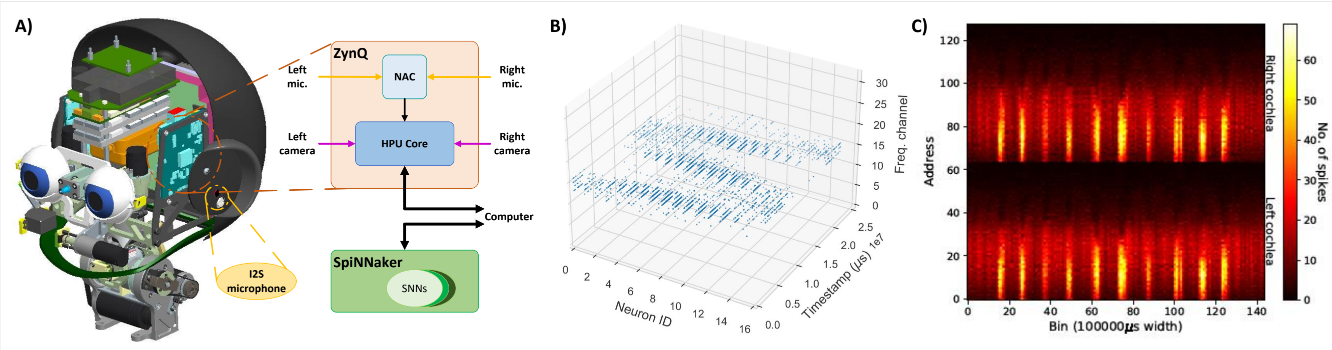

Figure 1: A) iCub-NAC integration diagram. B) Output events from the localization model. C) Sonogram from the keyword spotting.

source localization and is the basic building block for auditory attention [4].

Preliminary results of the NAC-iCub integration show that iCub was successfully able to localize a sound source which was shifted from left to right and viceversa in front of the iCub, as shown in Fig. 1 B). In addition, a dataset of spoken digits, shown in Fig. 1 C), was directly recorded from the iCub to train a SNN for performing keyword spotting. To the best of our knowledge, this is the first time that a digital neuromorphic auditory model has been integrated within a humanoid robot, and these results open the door for future implementations of sensory fusion models with learning algorithms to solve tasks such as the cocktail party problem on a fully neuromorphic real-time robotic platform, benefiting from the computational latency and low-power consumption nature of these systems.

ACKNOWLEDGMENTS

This work was supported by the Spanish grant MINDROB (PID2019-105556GB-C33/AEI/10.13039/501100011033), and SMALL (PCI2019-111841-2/AEI/10.1309/501100011033 from EU CHIST-ERA program). Part of this work was carried out during a research internship of D. G.-G. in the EDPR group (Italian Institute of Technology, Genoa (Italy)) supported by a Formación de Personal Investigador Scholarship from the Spanish Ministry of Education and Culture.

REFERENCES

[1] Jorge Dávila-Chacón et al. 2018. Enhanced robot speech recognition using biomimetic binaural sound source localization. *IEEE transactions on neural networks and learning systems* 30, 1 (2018), 138–150.

[2] Arren Glover et al. 2018. The event-driven software library for YARP—With algorithms and iCub applications. *Frontiers in Robotics and AI* 4 (2018), 73.

[3] Daniel Gutierrez-Galan et al. 2019. A neuromorphic approach of the sound source localization task in real-time embedded systems: work-in-progress. In *Proceedings of the International Conference on Embedded Software Companion*. 1–2.

[4] Dillon A Hambrook et al. 2017. A Bayesian computational basis for auditory selective attention using head rotation and the interaural time-difference cue. *PloS one* 12, 10 (2017).

[5] Bertrand Higy, Alessio Mereta, Giorgio Metta, and Leonardo Badino. 2018. Speech recognition for the icub platform. *Frontiers in Robotics and AI* 5 (2018), 10.

[6] Angel Jiménez-Fernández et al. 2017. A Binaural Neuromorphic Auditory Sensor for FPGA: A Spike Signal Processing Approach. *IEEE Trans. Neural Netw. Learning Syst.* 28, 4 (2017), 804–818.

[7] Lorenzo Natale et al. 2017. icub: The not-yet-finished story of building a robot child. *Science Robotics* 2, 13 (2017).

[8] Eustace Painkras et al. 2013. SpiNNaker: A 1-W 18-core system-on-chip for massively-parallel neural network simulation. *IEEE Journal of Solid-State Circuits* 48, 8 (2013), 1943–1953.

[9] Giulia Pasquale, Carlo Ciliberto, Francesca Odone, Lorenzo Rosasco, and Lorenzo Natale. 2019. Are we done with object recognition? The iCub robot's perspective. *Robotics and Autonomous Systems* 112 (2019), 260–281.

[10] Vadim Tikhanoff et al. 2010. Integration of speech and action in humanoid robots: iCub simulation experiments. *IEEE Transactions on Autonomous Mental Development* 3, 1 (2010), 17–29.