Live Demonstration: Real-time neuro-inspired sound source localization and tracking architecture applied to a robotic platform

F. Perez-Peña[1] E. Cerezuela-Escudero[2], Angel Jimenez-Fernandez[2] and Arturo Morgado-Estevez[1]

1. Applied Robotics Research Lab, Universidad de Cádiz, Faculty of Engineering, Puerto Real, Cadiz, Spain
2. Robotics and Technology of Computers Lab (RTC), University of Seville, ETSI Informática, Seville, Spain
Email: fernandoperez.pena@uca.es

Abstract— **This live demonstration presents a sound source localization and tracking system implemented with Spike Signal Processing (SSP) building blocks on FPGA devices. The system architecture is based on the ability of the mammalian auditory system to locate the direction of a sound in the horizontal plane using the interaural intensity difference. We used a binaural Neuromorphic Auditory Sensor to obtain spike rates similar to those generated by the inner hair cells of the human auditory system and the component that obtains the interaural intensity difference is inspired by the lateral superior olive. The spike stream that represents the interaural intensity difference is used to turn a robotic platform towards the sound source direction. The system was tested with pure tones (1-kHz, 2.5-kHz and 5-kHz sounds) with an average error of 2.32 degrees.**

I. INTRODUCTION

This live demo is based on reference [1]. It shows the hardware implementation of a sound localization and tracking system inspired by the mammalian auditory system. The NAS sensor used produces a biological cochlea-like output. This output is the stimulus for the processing system where the LSO model is implemented. The architecture proposed for the LSO which performs the subtraction between two input spike rates produces the Interaural Intensity Difference (IID). The IID auditory cue is used as the input for the spike-based actuation stage that tracks the sound.

The demo shows the system tested using 1 kHz, 2.5 kHz and 5 kHz pure tones. The maximum error obtained is less than five degrees. Furthermore, our system shows a high noise tolerance level when white noise is applied: in the worst condition, the average error is lower than ten degrees. The architecture presented in this demo is implemented by using low-cost commercial hardware devices such as FPGAs. The power consumption goes up to 58.33 mW in operation (29.7 mW from the NAS and 28.63 mW from the processing layer).

II. DEMOSTRATION SETUP

The experimental setup is shown in Fig. 1. It consists of a stimulus (sound source), a robotic platform (head), an auditory sensor and an actuation layer. The distance between the speaker and the head is 40 cm at different azimuthal angles (0º to 90º in steps of 15º). The microphones are on each side of the head (omnidirectional pick-up pattern). The head is placed on top of a platform driven by a DC motor with an encoder (Micromotor Ref. 2224R006SR plus gearhead Ref. 20/1 112:1 and encoder Ref. IE2-512 from Faulhaber). The NAS is implemented using a Virtex5 FPGA (XC5VFX70T) and it uses up to 99% of the total slices available. The FPGA is in a Xilinx development board (ML507) which includes the AC'97 audio codec. The NAS output is connected to the processing system using the AER protocol. The processing system is also implemented using a Virtex5 FPGA (XC5VFX30T), which uses up to 4% of the total slices available. The specifications of the microphone are: transducer principle based on back electret condenser element, the frequency response range is between 20 and 16,000Hz, the sensitive is -64dB ±3dB and the impedance is 1,000 Ohm.
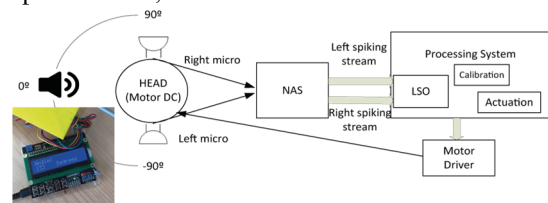

Fig 1. Set-up provided for the visitor experience.

III. VISITOR EXPERIENCE

Visitors will be able to interact with the demo by moving the audio source within the range (-90, 90) degrees and by modifying the tone played by the source: (1, 2.5 and 5) kHz. They will check if the head is able to follow the source and how accurate is the movement. There will be a small screen to show the ground truth and a laptop showing the current position reached by the head (Figure 1).

IV. REFERENCES

[1] E. Cerezuela-Escudero, F. Pérez-Peña et al., "Real-time neuro-inspired sound source localization and tracking architecture applied to a robotic platform," *Neurocomputing,* vol. 283C, pp. 129-139, Feb. 2018.