Parallel distribution of an inner hair cell and auditory nerve model for real-time application

Robert James, Jim Garside, Michael Hopkins, Luis A. Plana, Steve Temple, Simon Davidson, Steve Furber
The University of Manchester, Oxford Road, Manchester, M13 9PL UK
Email: robert.james@manchester.ac.uk

Abstract—This paper summarises recent efforts into implementing a model of the inner hair cell and auditory nerve on a neuromorphic hardware platform, the SpiNNaker machine. Such an implementation exploits the massive parallelism of the target architecture to obtain real-time modelling to a biologically realistic number of human auditory nerve fibres. The potential for incorporating this implementation into a full-scale digital real-time model of the human auditory pathway is then discussed.

I. THE AUDITORY PATHWAY

The mammalian ear extracts many useful features from a sound stimulus that have been essential in evolutionary survival. Such attempts to replicate these capabilities using various methods of frequency analysis and computer algorithms currently cannot match human performance in a range of hearing tasks [1] [2] [3]. It is believed that the brain’s unique stimulus encoding of spiking neuron action potentials, and the non-linear adaptive response of the inner ear, are contributing factors to what enables us to perform fast and efficient sound processing [4]. To achieve a realistic model of the mammalian ‘auditory pathway’ is the aim of this (and further) research. Using such a model we hope to gain a better understanding of the mechanisms in our auditory brain that explain human level hearing performance.

The auditory pathway begins with a sound pressure wave travelling into the outer ear and eventually displacing the Tympanic Membrane (TM) which separates the outer and middle ear. Inside the middle ear the TM connects to the cochlea via three ossicle bones to continue (and amplify) this displacement into the inner ear cochlea. The cochlea is a coiled, liquid filled organ that converts the TM displacement into a series of travelling waves along its distance, from base to apex. The frequency components of the sound stimulus dictate the location along the cochlea that will experience the most displacement along its Basilar Membrane (BM). High frequencies are absorbed at the basal regions and progressively lower frequencies are absorbed towards the apex of the cochlea. The cochlea is lined with many motion sensitive cells that detect the localised displacements of the BM, known as Inner Hair Cells (IHCs). Fig. 1 shows the structure of an IHC and the afferent Auditory Nerve (AN) fibres that innervate each IHC. The stereocilia ‘hairs’ at the top of the cell move according with the motion of the section of the BM the cell is closest to. When the stereocilia move in the axis of sensitivity they allow an influx of calcium ions (Ca$^{2+}$) to enter the cell body. This influx of Ca$^{2+}$ increases the likelihood of neurotransmitter release from the containing vesicles at the ‘active regions’ around the bottom of the IHC where AN fibres innervate the cell membrane. A release of vesicle neurotransmitter causes an action potential (spike) to occur in the corresponding AN fibre. Fig. 2 shows a diagram of an uncoiled cochlea; it illustrates how specific characteristics of a sound stimulus are represented by spiking activity in AN fibres.

The implementation featured in this work is based on a section of a model of the auditory pathway: the Inner Hair Cell and Auditory Nerve (IHC/AN). The method is based on the model described by Sumner et al. [5]. An overview of the algorithm and how this corresponds to believed biological processes is outlined in Section II-A.

To model this functionality on conventional computer hardware each IHC would have to be processed serially. This research effort uses a massively parallel neuromorphic hardware platform to perform the modelling in a biologically inspired fashion. A single frequency band BM channel will act as an input to an instance of the IHC/AN model, where many parallel instances are distributed across the same machine thus modelling all IHCs at once.
It is estimated that the human auditory brainstem is fed by approximately 30,000 AN fibres from each cochlea. There are around ten AN fibres that innervate each inner hair cell, each with slightly different characteristic responses defined by their ‘spontaneous rates’ [6]. We suggest that the number and variety of AN fibres that make up the cochlea output is vital in allowing the brain to represent sound features to the required spectrotemporal resolution in the auditory cortex.

II. MODELLING THE IHC/AN

A. Model Algorithm

The IHC/AN model can be split into four main sections. These correspond to a literal top-down approach to modelling the IHC: from the stereocilia to the auditory nerve synapse.

1) Receptor potential Here the displacement of the stereocilia is converted into a measure of how many ion channels open at the top of the IHC. A large number of open ion channels will cause an increase in the cell’s apical conductance. This conductance is used to calculate an accumulated membrane Receptor Potential (RP).

2) Calcium influx The calcium ion influx is modelled as multiple RP sensitive channels that transport calcium ions to the cell’s active regions. It is the variation in the conductance parameters of these individual channels that dictate the spontaneous firing rate of the corresponding AN fibre. The degree of spontaneous rate dictates the fibre response to the same stimuli. It is believed that this variation in responses is useful to how sounds are interpreted [7]. Fig. 3 shows different fibre type responses to the same increasing intensity stimuli. The channel Ca$^{2+}$ current is used to calculate the accumulated Ca$^{2+}$ concentration at each active region.

3) Vesicle release An increase in Ca$^{2+}$ concentration at the active regions of the cells increases the probability of vesicle release from the active region’s immediate vesicle store (ribbon synapse). The ultimate release of a vesicle is based on this probability determining a random process. A vesicle release will cause an ‘ejection’ of neurotransmitter into the synaptic cleft between the AN fibre and the IHC active region. The manufacture and replenishment of new vesicles to the cell active regions and cleft neurotransmitter re-uptake are included in the model as similar stochastic processes.

4) Auditory nerve action potential This model assumes the neurotransmitter in one released vesicle will trigger a spike in the AN fibre providing it is not already in its refractory (post firing) period.

Additional detailed descriptions of how the model algorithm is derived are presented in the literature [8] [5] [9].

B. Hardware implementation

The chosen hardware platform for this model is a SpiNNaker machine [10]. The SpiNNaker architecture allows for Real-Time (RT) execution of large scale spiking neural networks. Its hardware is made up of multiple ARM9 microprocessors (cores) which can run an individual application asynchronously with respect to their neighbouring cores. These processors can run any interrupt driven software model, not just spiking neuron models. The immediate ‘embarrassingly parallel’ nature of the IHC/AN model’s separate frequency band inputs allows for a large amount of parallel processing. The implementation presented here exploits this by running multiple instances of IHC/AN models as individual applications across parallel SpiNNaker cores. This provides scaling of an IHC/AN model to a biologically realistic number of channels with RT performance. Such performance would not be possible on conventional serial computation hardware. The low hardware clock rate of the SpiNNaker cores (200 MHz) produces much lower energy consumption than conventional processors whilst maintaining a high throughput due to the parallelisation of the computation.

C. Software conversion

The main efforts for this work have been converting the model algorithm to run on a SpiNNaker core. This involved converting the Matlab Auditory Periphery (MAP) [9] implementation of the model algorithm into an ARM9 executable.
written as an event-driven C application. The SpiNNaker implementation uses single precision software floating point representation of variables and uniquely seeded pseudo-random number generators.

D. Preprocessing

The input data to the IHC/AN model is generated by a preprocessing stage on a host computer running a full MAP simulation. These stages start at the outer and middle ear, through the tympanic membrane and into the cochlea. The cochlea model is performed using a Dual Resonance Non-Linear filter-bank (DRNL) [11]. This models the cochlea as a parallel filter-bank where each filter ‘channel’ output represents the displacement of a segment of the BM along the cochlea. These DRNL outputs provide the parallel inputs to the IHC/AN models.

E. Results

The execution times for a 454 ms ‘A’ vowel stimulus are shown in Fig. 4. The SpiNNaker implementation processing time is, on average, 21.7 µs per sample for any number channel IHC/AN model. This will allow for a RT processing of an input stimulus at a sampling frequency of 44.1 kHz. Fig. 4 shows the desktop computer processing time increases with an approximate linear relationship with number of model channels.

The model outputs from the ‘A’ vowel stimulus are shown in Fig. 5b and were obtained from running a 768 channel (1536 fibre) version of the IHC/AN model across 48 SpiNNaker chips. Both plots show the occurrences of spikes along the output AN fibres. Each model instance produces two AN fibres, one of low (LSR) and the other a high (HSR) spontaneous rate. The limit of two fibre outputs per model is to achieve RT performance on SpiNNaker. This is fewer than the ten observed fibres that innervate a single IHC however the SpiNNaker simulation is flexible enough to allow for multiple instances of the model to use the same input, thus allowing for a reconfigurable number of fibres of either type to be allocated.

Due to the stochastic nature of the model outputs any comparisons of individual runs will show variation. Therefore to perform a credible comparison between both model implementations Post Stimulus Time Histogram (PSTH) plots of the model outputs have been generated. The results shown in Fig. 6 show the time varying AN spike rates across 1 ms windows to a 6.9 kHz sinusoidal 68 dB SPL stimulus, first in Fig. 6a from physiological data gathered by Westerman and Smith [12] and then from both model implementations in Fig. 6b. These results show both implementations produce a biologically similar response consisting of a pre-stimulus response of approximately 50 sp/s, followed by a peak response at stimulus onset at around 800 sp/s, decaying to an adapted rate in the region of 170 sp/s. Finally at stimulus removal rates significantly drop during an offset period before returning to spontaneous firing rates of approximately 50 sp/s.

III. Future Work

A. Larger simulations

Following this investigation we intend to increase simulations to a human number of AN fibres (30,000) on a SpiNNaker machine of 15,000 cores. Such an implementation can be accommodated on the available machine of 500,000 cores and would consume power of approximately 1.5 kW.

B. Full Auditory Pathway Model

We plan to incorporate this model into a large scale full simulation of the auditory pathway. The major benefit of having the IHC/AN model output already within the SpiNNaker fabric is it allows for interfacing with subsequent auditory brainstem, midbrain and cortical neuron models running on the same machine. For this publication the IHC/AN implementation was executed and tested in isolation using preprocessed inputs. More recent efforts have gone into integrating a complete pipeline version of the MAP model on SpiNNaker hardware. The unique multi-cast routing protocol used on SpiNNaker provides a scalable solution with
the ability to model the complex network of ascending and descending projections in the auditory pathway and gives the potential for RT, interactive interfaces.

IV. CONCLUSION

This work has shown the feasibility of implementing a human IHC/AN model to a biological scale without compromising processing performance. Using a neuromorphic platform for this and later processing of the auditory pathway human perception may be better understood. This digital implementation has flexibility to be refined according to future developments. The output of this model will interface directly with other spiking neural processors across the SpiNNaker fabric.

ACKNOWLEDGEMENT

We thank Professor Ray Meddis for his support and advice, EPSRC EP/D07908X/1 and EP/G015740/1, ERC 320689 and HBP FP7-604102 for funding the SpiNNaker project.

REFERENCES


