Hearing Thinking

Jane Grant, John Matthias, Tim Hodgson and Eduardo Miranda
Interdisciplinary Centre for Computer Music Research (ICCMR)
Faculty of Arts and Faculty of Technology, University of Plymouth,
Drake Circus, Plymouth PL4 8AA
J1grant@plymouth.ac.uk john.matthias@plymouth.ac.uk tim.hodgson@plymouth.ac.uk
eduardo.miranda@plymouth.ac.uk

Abstract. This paper describes early experiments, which attempt to reconfigure the sound of a breath using a network of artificial spiking cortical neurons. The connectivity of the network evolves according to a spike timing dependent plasticity algorithm and the instrument triggers grains of sound from recordings of the breath when any of the neurons fire.

Keywords: Cortical Neurons, Granular Synthesis, Synaptic Plasticity

1 Introduction

When a person sees, light enters the eyes. The information is translated into electrical signals via sensory neurons and is processed in the brain via cortical neurons. The information is then interpreted by the brain, and any action taken will be via motor neurons which will result in external motor action.

When a person hears, sound enters the ears. The information is translated into electrical signals via sensory neurons and is processed in the brain via cortical neurons. The information is interpreted by the brain, and any action taken will be via motor neurons which will result in external motor action.

In terms of vision, we can define the process involving the seeing, the sensory and cortical processing and the interpretation as ‘looking’. In terms of audition, we can define the sensory and cortical processing and the interpretation as ‘listening’
This project began with an specific interest in the relationship between our sensual perception of the external world (for example, looking and hearing) and the externalization of that perception in relation to the idea of ‘self’. The specific externalization (action) focused on is the voice and more precisely the breath, the beginning of vocalization. Daniel C. Dennett [1] proposes the notion of ‘self’ as a construction; that ‘the self’ is a ‘naive boundary between “me” and “the outside world”’ and suggests that ‘even such a simple self is not a concrete thing but just an abstraction, a principle of organisation. Moreover the boundaries of a biological self are porous and indefinite.’ In the field of phenomenology, David Wood [2] argues that such boundaries are interchangeable as states, that ‘a boundary is not a thing, but a cluster of procedures for the management of otherness’. The central artistic aim is to

affect a rupturing of the boundaries between the sensed and the action and explore the effects of removing the sensing and sensory part of the 'self'. In order to examine the effect of such a rupturing, we have merged an externalised action (the recorded sound of a breath) with a sonic instrument, which creates sound from a model of cortical neuronal firing patterns known as the Neurogranular Sampler [3].

The Neurogranular sampler works by triggering grains of sound (typically in a range of duration of 10 milli seconds (ms) -50ms) taken from a recorded sample when any one of an isolated network of artificial cortical neurons 'fires'. The resulting sound therefore consists of short bursts of the original sample triggered by the cortical neurons. It is a sonification of the cortical firing patterns. In this work, the sound of both voice and breath are recorded then reconfigured via the Neurogranular sampler in order to merge the voice or breath with the patterns and rhythms occurring in the neuronal network; we 'hear' neurons firing but within the recognised 'sound frame' of the human breath. The voice, but more particularly the breath itself belongs to an integral internal to external process and is one that we all recognise readily. There is a transition from internal to external, a process which might result in the listener hearing both the alien and the familiar simultaneously, existing at a liminal threshold.

2 The Neurogranular Sampler

The Neurogranular sampler [3] utilizes the recently developed model of a spiking neural network of Izhikevich et. al. [4]. The model contains N cortical neurons, each of which are described by two dimensionless variables v_i and u_i, where v_i represents the membrane potential of the ith neuron and u_i represents a membrane recovery variable, which provides negative feedback to v_i. The system is then described by the following coupled ordinary (nonlinear) differential equations:

$$\frac{dv_i}{dt} = 0.04v_i^2 + 5v_i + (140 - u_i) + I_i \tag{1}$$

$$\frac{du_i}{dt} = a(bv_i - u_i) \tag{2}$$

with the following auxiliary condition after spike resetting; if $v_i \geq 30$ millivolts then $v_i \rightarrow c$ and $u_i \rightarrow (u_i + d)$. Essentially, this means that when a neuron receives a spike input then its membrane potential is immediately reset.

The neurons are coupled to one another through a matrix of synaptic connection weights, which are contained in the matrix $S = (s_{ij})$. Synaptic currents or injected dc-currents (currents which come from either other neurons or from sensory information) are encompassed within variable I and a, b, c and d are parameters which describe the individual neuron characteristics. Essentially, different values of these parameters produce different individual intrinsic neuron firing patterns such that complex spiking, bursting or chattering of cortical and thalamic neurons can be simulated [3].

The methodology of sonification we use for the Izhikevich model uses the time of firing as the main sound event. All of the parameters in the model are used as controllers for the time of firing for each neuron. Every time a neuron fires, a dot is placed in a sound event file. We use these events as temporal points at which sounds are generated. The dynamics of the spiking network are determined and controlled by varying the number/type of neurons and the geometry of the connectivity In the original work of Miranda and Matthias [3], the elements of the matrix of connections, *S* were updated using the simple Izhikevich algorithm [4]. In the simple algorithm, elements of *S* are fixed and added to the postsynaptic membrane potential instantaneously following a firing event. In the current implementation, the elements of the matrix *S* of connections are updated either using the simple model or according to a Spike Timing Dependent Plasticity (STDP) algorithm [5]. In this algorithm, connections for which pre-synaptic spikes cause post-synaptic firing are potentiated and those for which pre-synaptic spikes arrive after post-synaptic firing has occurred are depressed. We also have included axonal conduction delays [6], which are random numbers between 1ms and the maximal conduction delay parameter, which in our implementation is of the order of 10-20ms.
The neuronal network remains isolated to any sensory information and is unstimulated from outside the network. It is driven by noise such that the initial current and voltage parameters within a proportion of neurons will be of a high enough value to drive them to fire. Figure 1 shows a typical raster plot in a simulation of 1000 neurons. We can see that the firing response is highly correlated and that the firing forms a coherent pulse across the network.

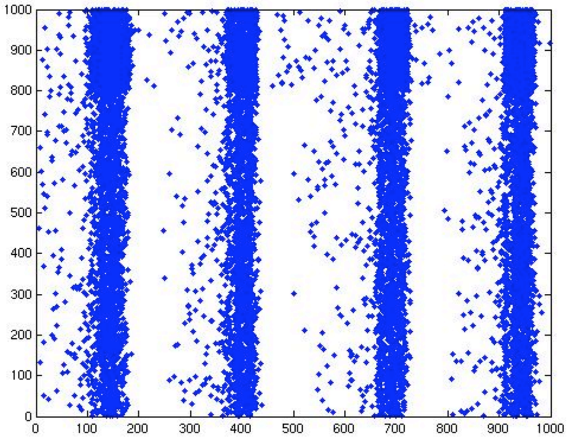

Fig. 1. A graph showing neuron number (y-axis) plotted against time (milli seconds, x-axis) in a simulation from the Izhikevich model including spike timing dependent plasticity and axonal conduction delays. The maximun axonal conduction delay parameter in this simulation is 20ms.

2.1 Sonification

In our methodology, short segments (or "sound grains") taken from sound files are triggered when any of the neurons fire. The algorithm by which the system works is straightforward: When a neuron in the network fires at time *t*, a sound grain of predetermined length (between 10-50ms) and amplitude is taken from the recorded sample of sound and played back. The sound grain is convoluted within a Hanning envelope (7). Synchronized firing of neurons sound like a pulse, whilst networks containing only a few neurons produce sparse rhythmic sequences. The system therefore has a very wide variety of temporal patterns and behaviours, which can be controlled according to the parameters of the model. One can control the parameters *a*, *b*, *c* and *d*, which determine the intrinsic properties of the neurons and one can control the number and type of neurons. Generally speaking, increasing the number of neurons in the model means more firing and therefore more sonic texture, although when the solutions exhibit synchronous behaviour, increasing the number of neurons tends to lower the frequency of the collective response. It is interesting in itself that such random (noisy) inputs can produce synchronous rhythms, a well-understood phenomenon within the dynamical systems community. Figure 2 shows the output amplitude from the Neurogranular Sampler against time in an early simulation which consists of 4 excitatory Neurons and 1 inhibitory neuron with an axonal delay of up to 10ms including STDP.

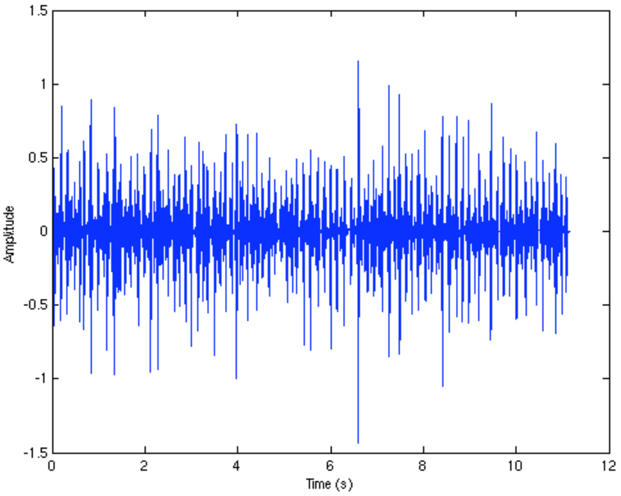

Fig. 2. A simulation of the Neurogranular Sampler which shows the amplitude of signal plotted against time in seconds for an artificial network using the Izhikevich model with 4

excitatory neurons and 1 inhibitory neuron all of regular spiking type including an axonal delay of up to 10ms and Spike-timing dependent plasticity. The initial voltage and current initial parameters were taken as random matrices. The external sound in this example is taken from a single note played on a harmonium.

The STDP algorithm along with the axonal delays encourage particular pathways in the network to become established, which lead to more events and more regular frequency of neuronal firing. In the absence of sensory input to the network, these pathways have a transient lifetime.

3 Methodolgy

The process of experimentation for this work initially consisted of setting up a user interface for the Neurogranular Sampler. This was written in MATLAB and enabled the user to choose the number of neurons in the network, individual neuronal firing characteristics, an initial matrix for the initial current, I input to each neuron and the sound file for sampling. The user has a choice whether to implement a dynamic matrix of connections S via the STDP rule, or to use the simple static initially configured matrix of connections. Experiments with the interface consisted of inputting a sequence of breaths and vocal sounds into the sampler. The parameters on the interface were chosen in order to achieve a rupturing effect such that the sound of the (external) breath is reconfigured to enable the listener to hear the external reconfigured within the context of the (internal) firing.

The initial experiments of Miranda and Matthias [3] use very short grains, the result being a direct rhythmical translation of the simple neuronal firing patterns. In addition, the effect of using very short grains renders the result rather decoupled from the initial recording. In this work, in order to affect the rupturing between the breath and the processing, we used much longer grains than had previously been used. In fact the possible duration of grain sizes has been greatly increased, to a point where they cannot really be considered 'grains' any more. The resultant layering of multiple samples gives rise to some intriguing sonic effects: When using the larger grain lengths, we experienced a real lengthening and spatial aspect to the experimental outcomes. We also decided to depart from the original sampler design, in which grains were taken at random from any point in the input sound file. We felt that this removed too much of the structural character from the source material. Instead, the user can define a 'grain window', limiting how far in time from its original position any given grain can move. This may be anything from 1ms to the entire length of the output file, giving the user great flexibility in determining how radically the Neurogranular sampler transforms the input. 5n this implementation, the breath or voice appears to be drawn out, but simultaneously held together by a series of ‘strands’ of information. These strands merge to form a singular long breath ruptured at times by waves winding through the sample. Figure 3 shows the output from a simulation from a recorded breath using a neuronal network including 6 excitatory and 6 inhibitory neurons. The grain size was allowed to vary between 695 and 4352 samples for which the sampling rate was 44.1kHz.

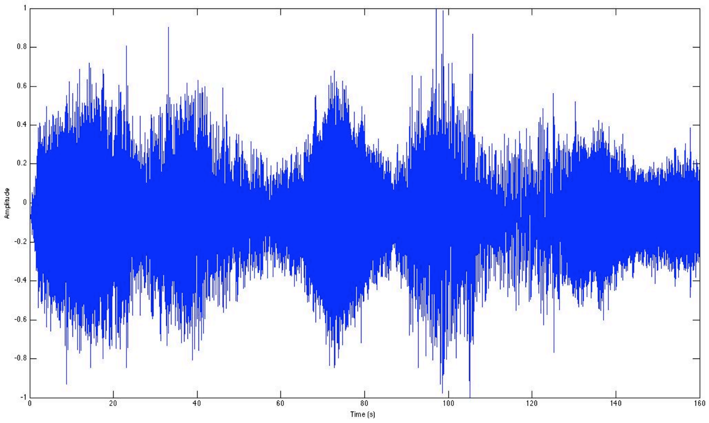

Fig. 3. A simulation of the Neurogranular Sampler which shows the amplitude of signal plotted against time in seconds for an artificial network using the Izhikevich model with 6 excitatory neurons and 6 inhibitory neuron all of regular spiking type including an axonal delay of up to 20ms and Spike-timing dependent plasticity. The recorded sound was a breath, which lasted for approximately 10seconds. The grain size was allowed to vary in duration between 695 and 4352 samples. The sampling rate was 44.1kHz.

4 Discussion

In this work, there is no sensory information input to the network of cortical neurons. The ‘sensory’ information exists in the recorded sample, which is retriggered and reconfigured according to the firing patterns of an isolated tiny cortex. There is no ‘hearing’ or ‘seeing’ but a reconfiguring of a sound already heard. The sounds emulate the strangeness of what we might imagine the merging of the voice with a neural structure could sound like, the more interesting results causing a tension between what we perceive to be an internalized sound but resonating with an external spatial aspect. In further work, we intend to explore how the system might respond when sensory information is directly input to the neuronal network, integrating the sensory/cortical interaction. We also intend to explore the resulting dynamic evolution of the synaptic connections and firing pathways in the light of this sensory input, and additionally, produce a more interactive environment in which the user can work with the network in real time, enabling them to adapt the input according to the results of the processing.

As discussed at the beginning of this paper, one of the starting points in this work is the philosophical premise that the idea of ‘self’ is a construction, that ‘the self’ has no concrete boundaries. Daniel C. Dennett argues [1] that the boundaries of a biological self are ‘porous and indefinite’. The experiments outlined in this paper will eventually form the audio part of a sound and video installation called ‘Threshold’, by Jane Grant. This work is concerned ideas outlined here and also with the sense of space both infinitesimal and infinite, which the breath or voice inhabits. Gaston Bachelard in his book, ‘The Poetics of Space’ [8] suggests that there lies an infinity within the pronunciation of the word ‘vast’, that in saying the word it ‘teaches us to breathe with the air that rests on the horizon…’[i]

The final work for this installation will send waves of sound undoing the metaphorical architecture of the brain and dispersing it towards the horizons.

References

1. Dennett, D.C.:Consciousness Explained, Penguin, (1991)
2. Wood, D.: Time shelters: an essay in the poetics of time. In: *Time and the Instant: essays in the physics and philosophy of time*. ed. Durie, R. Manchester. Clinamen Press (2000)
3. Miranda, E.R. and Matthias, J.R. : Granular Sampling using a pulse coupled network of spiking neurons, Proceedings of EvoWorkshops 2005, Lecture notes on Computer Science 3449 pp539-544 Berlin: Springer-Verlag
4. Izhikevich, E.M.: A simple model of a spiking neuron, IEEE Transactions on Neural Networks 14:1469
5. Song, S., Miller, K.D., Abbott, L.F., Competitive Hebbian Learning through spike-timing dependent synaptic plasticity Nat Neuroscience 3: 919-926 (2002)
6. Inzhikevich, E.M., Gally, J.A., and Edelman, G.M. Spike Timing Dynamics of Neuronal Groups, Cerebral Cortex 14:933-944 (2004)
7. Abramovitz, M. and Stegun, I.: Handbook of Mathematical Functions New York (USA) Dover (1972)
8. Bachelard, G.: The Poetics of Space, Beacon Press, (1964)

---

[i] Gaston Bachelard, The Poetics of Space, Beacon Press, 1964