Evaluation of a GPU-based auditory functional model with physiological data

Kazuyuki Sano, Kenji Ozawa*, Masanori Morise, Yu Sato and Sohei Chimoto

Graduate Faculty of Interdisciplinary Research, University of Yamanashi,
4-3-11 Takeda, Kofu, 400-8511 Japan

(Received 30 September 2018, Accepted for publication 30 April 2019)

Keywords: Auditory functional model, GPU, Physiological data, FM tone, Brain-machine interface

PACS number: 43.64.Bt, 43.64.Sj [doi:10.1250/ast.41.373]

1. Introduction

Cochlear implantations have been applied to more than 300,000 people worldwide. However, an auditory prosthesis to restore hearing to individuals with total hearing loss has not yet been developed. Notably, some studies have examined electrical stimulation to the auditory cortex (A1) [1,2]. The ultimate auditory prosthesis will be a brain-machine interface (BMI) in which electrical stimulation, similar to the responses of healthy A1 neurons to a sound, is fed into the A1 neurons of the impaired listeners. In order to produce such responses to a sound, we previously developed a functional model of the auditory pathway [3–5]. However, our model required an extended period of time to produce responses because it was implemented on a central processing unit (CPU).

We thus rebuilt our model with a graphics processing unit (GPU) to realize parallel and quick processing, in order to simulate the function of the auditory neural pathway. In this report, the performance of the model is evaluated by using physiological data with respect to responses to frequency-modulated (FM) tones.

2. Implementation of the model with a GPU

2.1. Construction of the model

Figure 1 depicts a block diagram of our functional model, consisting of seven stages: the basilar membrane (BM), inner hair cell (IHC), primary auditory nerve (AN), ventral cochlear nucleus (VCN), inferior colliculus (IC), medial geniculate body (MGB), and A1. From the fourth stage (VCN), because each channel has 64 neural units, the auditory pathway is represented with 4,096 neural units in parallel. Subsequently, frequency weightings are introduced between stages to simulate lateral inhibition in auditory neural networks. The model was installed by using a GPU (NVIDIA, Tesla k20c) with the C language programming environment (Microsoft, Visual Studio 2010).

Sixty-four-channel Gammatone filtering [6] in the BM stage is processed with a fast Fourier transform library (NVIDIA, CUFFT). The IHC stage is based on the Meddis model [7]. The AN, VCN, IC, MGB, and A1 stages are created based on Maki *et al.* [8–10]. The outputs from BM and IHC stages are waveforms, while those from the subsequent stages are spike trains that emulate neural firing.

To realize parallel processing from the VCN, each streaming multi-processor (SM) in the GPU is assigned to one of the 64 frequency channels; each streaming processor (SP) in a SM is assigned to one of the 64 units (Fig. 2). As shown in Fig. 2, each neuron is assigned to a *thread*; the set of 64 threads for each frequency channel constitutes a *block* in GPU programming. Therefore, there are 64 neurons with the same characteristic frequency (CF), each of which can simulate the different responses of a sustained-ON or phasic type [11] by properly setting values of parameters in the models [8–10]. Though these two types are typical neurons in the A1 of awake cats [11], they are only examples to estimate the potential of the implementation and do not cover all the auditory processing.

The number 64 was chosen for frequency channels because it was appropriate for GPU programming. However, it is difficult to argue whether the number 64 is sufficient or not at the current stage from the frequency resolution perspective although it is larger than the number of channels used in cochlear implantation, such as 24. The proposed system will give a workbench to investigate the issue of enough frequency resolution for an auditory prosthesis. Also, 64 neurons in one frequency channel should be argued to be enough or not as an approximation of the auditory pathway to represent different types of cells in the auditory system.

2.2. Behavior of the model

The aforementioned processing from the BM through A1 stages is conducted sequentially for data in every 50-ms temporal segment where the sampling frequency is 48 kHz. Here, task-elapsed time is measured for the processing of all 4,096 neurons and it is independent of the simulated cell types. The task-elapsed time for one segment is 1.5 s while it was 144.7 s with our previous CPU (Intel, Xeon E5-2630) implementation. Although the execution speed is approximately 100 times faster than our previous CPU implementation, real-time processing is not available. This challenge will be surpassed by using the latest GPU (e.g., NVIDIA, GeForce GTX 1080 Ti) and *grid* processing in GPU programming. In the *grid* processing, as *blocks* are declared multi-dimensionally, the number of parallel processing increases and therefore faster processing is expected. Moreover, if the target sound is a speech signal, reducing the sampling frequency to 8 kHz will be effective for reduction of the quantity of data per segment.

*e-mail: ozawa@yamanashi.ac.jp

Fig. 1 Block diagram of the auditory functional model.

Fig. 2 Implementation of 4,096 units in parallel, where SM is a streaming multi-processor and SP is a streaming processor. These 4,096 units are embedded in each stage of VCN, IC, MGB, and A1 of Fig. 1.

3. Evaluation of the model

The properties of the model were evaluated by comparing model output with physiological data. These data were observed from the A1 neurons of awake cats as the responses to FM sweeps [11]. The stimuli used were upward or downward sweeps between 0.2 and 8.0 kHz with various sweep speeds: 4, 8, 16, 32, 64, or 128 Hz/ms.

Dotted lines in Fig. 3 show the data observed from a sustained-ON-type cell (CF: 3.2 kHz) denoted in a 1-ms bin peri-stimulus time histogram (PSTH). The PSTH was drawn as the average of 10 stimulus presentations with smoothing by a Gaussian function of 5-ms SD [11].

The stimuli used in the physiological experiment were fed into the model; PSTHs were drawn in the same manner. Table 1 shows the parameter values used in the model to simulate a sustained-ON-type cell. Solid lines in Fig. 3 denote the model outputs. Irrespective of sweep speeds, the model outputs show good approximation of the physiological data for upward FM sweeps. Similar results were obtained for downward FM sweeps.

Evaluations were also conducted for phasic-type cells. A model with appropriate parameter values could sufficiently simulate the physiological data. Thus, we conclude that the model can represent different types of A1 cells and may be appropriate for use as an auditory BMI. Further examinations are required for a variety of sound stimuli, including speech.

Acknowledgment

The authors thank Mr. Shohei Higashimoto of the

Fig. 3 PSTH of A1 neural responses of a sustained-ON-type cell to upward FM tones (Solid line: model outputs, Dotted line: physiological data). Bold bars below the temporal axes represent the duration of each stimulus presentation.

Table 1 Parameter values used to simulate a sustained-ON-type cell in A1 neurons.

Stage	VCN	IC	MGB	A1
a_i	1.3×10^5	1.5×10^5	0.8×10^5	0.2×10^5
τ_i	6.66×10^{-6}	1.5×10^{-5}	8.0×10^{-6}	3.0×10^{-5}
μ_c	0.0008	0.003	0.005	0.01
σ_c	0.00005	0.00015	0.0002	0.005
μ_r	0.0008	0.0014	0.0015	0.0095
σ_r	$0.025 \times \mu_r$	$0.1 \times \mu_r$	$0.25 \times \mu_r$	$0.25 \times \mu_r$
β	7.0	3.0	0.5	0.5
α	1.0	0.2	0.13	0.22

University of Yamanashi for his contributions in GPU programming.

References

- [1] W. H. Doherty, S. S. Stensaas, M. G. Mladejovsky and J. B. Smith, "A prosthesis for the deaf based on cortical stimulation," *Ann. Otol.*, **82**, 445–463 (1973).
- [2] J. Wang, Y. Liu, L. Qin, S. Chimoto, K. Nakamoto and Y. Sato, "Chronic microstimulation of cat auditory cortex effective to evoke detection behaviors," *Neuroscience*, **206**, 81–88 (2012).
- [3] T. Goda, T. Kishita, K. Ozawa, L. Qin and Y. Sato, "Modelling responses of the primary auditory cortex to frequency-

- modulated tones in awake cats,” *Proc. 19th IASTED Int. Conf. Appl. Sim. and Model. (ASM 2011)*, pp. 161–167 (2011).
- [4] T. Goda, K. Ozawa, Y. Kinoshita, Y. Sato and S. Chimoto, “Functional model of the level sensitivity of a tonic cell in the primary auditory cortex of awake cats,” *J. Acoust. Soc. Jpn. (J)*, **67**, 395–401 (2011) (in Japanese).
 - [5] K. Ozawa, T. Goda, L. Qin and Y. Sato, “Functional model to replicate the representations of human vowels in the primary auditory cortex,” *Proc. 20th IASTED Int. Conf. Appl. Sim. and Model. (ASM 2012)*, pp. 72–77 (2012).
 - [6] R. D. Patterson, M. Allerhand and C. Giguere, “Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform,” *J. Acoust. Soc. Am.*, **98**, 1890–1894 (1995).
 - [7] R. Meddis, “Simulation of auditory-neural transduction: Further studies,” *J. Acoust. Soc. Am.*, **83**, 1056–1063 (1988).
 - [8] K. Maki, K. Hirota and M. Akagi, “A functional model of the auditory peripheral system: Responses to simple and complex stimuli,” *Proc. NATO/ASI Computational Hearing Conf.*, pp. 13–18 (1998).
 - [9] K. Maki, M. Akagi and K. Hirota, “A computational model for temporal response patterns of neurons in the inferior colliculus,” *J. Acoust. Soc. Jpn. (J)*, **60**, 304–313 (2004) (in Japanese).
 - [10] K. Maki, M. Akagi and K. Hirota, “Responses to amplitude-modulated-tones from a ventral cochlear nucleus cell model,” *J. Acoust. Soc. Jpn. (J)*, **59**, 13–22 (2003) (in Japanese).
 - [11] L. Qin, J. Y. Wang and Y. Sato, “Heterogeneous neuronal responses to frequency-modulated tones in the primary auditory cortex of awake cats,” *J. Neurophysiol.*, **100**, 1622–1634 (2008).