

# Electronic Cochlea: CAR-FAC Model on FPGA

Ying Xu, Chetan Singh Thakur, Ram Kuber Singh, Runchun Wang, Jonathan Tapson, André van Schaik  
The MARCS Institute, Western Sydney University, Kingswood, NSW 2751, Australia  
[ying.xu@westernsydney.edu.au](mailto:ying.xu@westernsydney.edu.au)

**Abstract**—In the mammalian auditory pathway, the cochlea is the first stage of processing where the incoming sound signal is mapped into various frequency channels. This multi-channel output contains subtle spectrotemporal features essential in sound source identification and segregation. These features have a large dynamic range. They have a large gain and sharp tuning at low sound levels, and a low gain and broad tuning at high sound levels. These properties are inherent in the ‘Cascade of Asymmetric Resonators with Fast-Acting Compression (CAR-FAC) model of the cochlea. The CAR segment of the model simulates the basilar membrane’s response to sound while the FAC segment simulates hair cells and sub-cortical responses resulting in a neural activity pattern necessary for sound perception analysis. We have previously implemented the CAR model on an FPGA, and in this paper, we implement the complete CAR-FAC model with 100 sections on an FPGA for a sample rate of 32 kHz. The FAC part includes outer and inner hair cell algorithms incorporating nonlinearity as well as an automatic gain control for modulating basilar membrane output with respect to input sound levels. We have implemented our design using time-multiplexing approach, where hardware resources of one cochlea section are reused for all the 100 sections of the model.

**Keywords**—*Electronic cochlea; outer hair cell; inner hair cell; Neuromorphic engineering*

## I. INTRODUCTION

The auditory pathway is a complex sensory system that processes sounds in mammals and shapes our experience of the environment we are immersed in. The auditory pathway can be segmented and modeled into individual sub-systems. Models of such systems are beneficial in offering insights into the sophisticated mechanisms of the auditory system and to validate theories of audition [1]. They are used for developing hearing prosthetics and are potentially utilized in auditory research as well as in sound and music processing fields. One such model that efficiently [2] describes the dynamics of the first part of the auditory pathway, the cochlea, is the ‘Cascade of Asymmetric Resonators with Fast Acting Compression’ (CAR-FAC) [3]. The CAR-FAC model provides a good description of the auditory periphery in many ways. First, the CAR structure in the model provides an archetypal data fit on human tone detection in masking noise with fewer parameters than other models, such as the roex and gammachirp models [2]. Second, the FAC part includes dynamic nonlinear effects present in human hearing, such as fast wide-dynamic-range compression. Finally, the FAC part is realized via adapting the radius of the poles and zeros of the CAR in the z-plane. The fact that it is able to implement this in digital filters in an online fashion is a major reason for using this model. We will

describe briefly the sound processing capability of a mammalian cochlea in this section and its implementation in the next section. Thereafter, an FPGA design of the CAR-FAC model is elaborated.

In mammalian species, sounds are channeled to the inner ear via the outer and the middle ear. The middle ear serves as an impedance matching transformer that matches low impedance at the outer ear to the high impedance of the fluid-filled inner ear and thus transmits sound across two different mediums [4]. In the cochlea, sound mechanically vibrates a trapezoid-shaped tissue known as the basilar membrane (BM) [5]. Inner hair cells (IHC) placed on top and along the BM transduce mechanical motion to electrical spikes in the auditory nerve that are sent upstream to the cochlear nucleus and beyond for auditory sensory and perception processing [6]. Feedback connections from the brain are sent down to outer hair cells (OHC) placed along the BM to modulate the vibrational amplitudes of the travelling waves to be within the range of safe audition [7], [8].

Computational models of the cochlea executed on serial computers often experience bottlenecks in terms of processing time due to their complexities [9]. Appending such cochlea models with sub-cortical and cortical models of the brain, where the levels of sophistication increase manifold, only exacerbates this issue. A potential solution lies in analog VLSI implementation of the cochlea models where fast, low-powered and parallelized processing capable of handling noisy inputs is made possible [10]–[14]. However, the process of developing such analog VLSI devices takes considerable time and also their performance deteriorates with time due to transistor mismatch [15]. An FPGA solution offers a shorter development time, fast processing speed, and transfer characteristics that match a software model exactly. Furthermore, the improvement of FPGA technology allows large-scale neuromorphic systems to be accommodated on a single chip [16].

We have previously implemented the CAR model describing the sound-induced travelling waves phenomena of the basilar membrane on an FPGA [17], [18]. In this paper, we will elaborate on the implementation of the FAC functions of the CAR-FAC model on an FPGA. The features include IHC and OHC algorithm blocks as well as an automatic gain control (AGC) defining a feedback mechanism from sub-cortical regions of the auditory pathway in the brain that collectively modulate the travelling waves in the BM.

## II. CAR-FAC MODEL

The CAR model displayed in Fig. 1(a) describes the transformation of input sounds to the equivalent of a mechanically induced travelling wave on the BM. It utilizes a cascade of parallel filter banks (Fig. 1(a)), similar to transmission-line models, to map spectral contents of input sounds to various channels. Each channel has a skewed bell-shaped curve frequency response with a center frequency known as best frequency (BF). The output of each channel of the CAR model is defined by a two-pole-two-zero transfer function known as a pole-zero filter cascade (PZFC), visually described in Fig. 1(a).

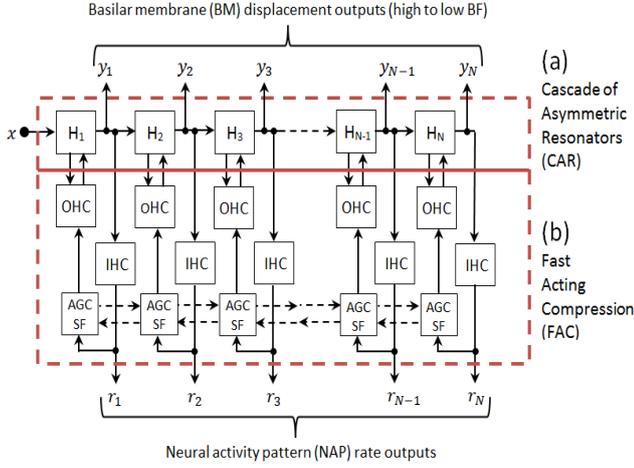


Fig. 1. CAR-FAC model.

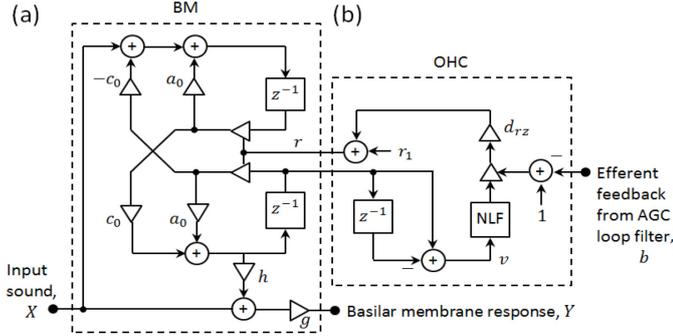


Fig. 2. Basilar membrane (BM) and outer hair cell (OHC) segments.

The CAR model coefficient,  $r$ , is coupled with an OHC model which incorporates an instantaneous damping mechanism giving a fast acting compression (FAC) affecting the BM travelling wave amplitudes. The FAC effect influences the variation of poles and zeros in the CAR model transfer function through the change in  $r$ . Thus, the radius of cosine and sine pole angles,  $a_0$  and  $c_0$ , in the  $z$ -plane are controlled, thereby affecting the BM displacement. These FAC effects in the OHC are computed using the BM velocity,  $v$ , as part of a local nonlinear function,  $NLF$ :

$$NLF(v) = 1/(1+(0.1v+0.04))^2 \quad (1)$$

This nonlinearity is scaled by gains  $d_{rz}$  and  $1-b$  (Fig. 1(b)). The former signifies the rate at which the  $NLF$  output influences the pole radius,  $r$ , while the latter depends on the feedback of the AGC loop-filter described below (Fig. 4).

The output of the CAR model,  $y$ , is conditioned by the IHC feedback, which is illustrated in Fig. 3. The first stage of the IHC is a high-pass filter (HPF) that suppresses frequencies below 20Hz. The next stage is an NLF, where the signal is half-wave rectified (HWR) and conditioned by a sigmoid-like function:

$$z = HWR(x + 0.175) \quad (2)$$

$$g = z^3 / (z^3 + z^2 + 0.1) \quad (3)$$

where,  $x$  is the high-pass filtered input;  $HWR$  is a half-wave rectification process; and  $g$  is the output of the IHC nonlinear function. This nonlinearity represents the motion of hair-like cilia bundle atop the IHC affected by the mechanical travelling waves of the BM. An AGC loop adapts the output of the IHC nonlinearity as a function of time. Two low-pass filters (LPFs) smoothen the output, and generate neural activity patterns (NAPs) of the auditory nerve.

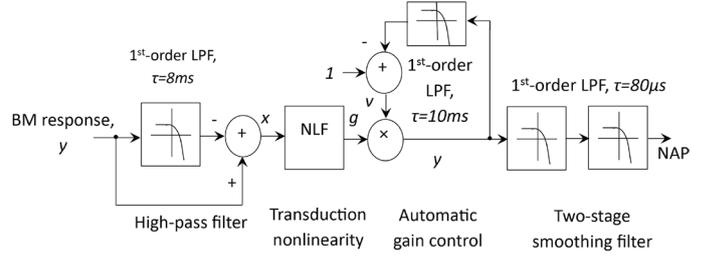


Fig. 3. Inner hair cell (IHC) segment.

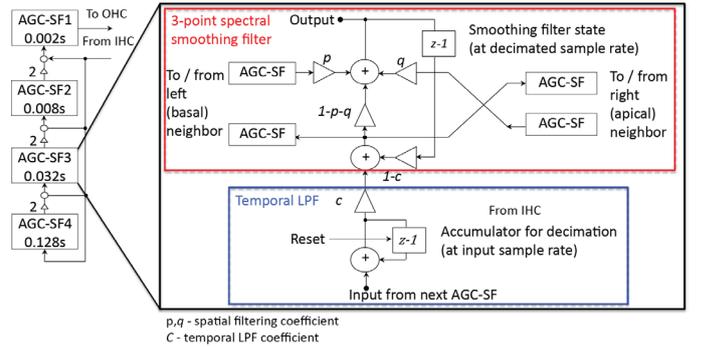


Fig. 4. Automatic gain control (AGC) with spectrotemporal smoothing filters (SF). Four stages of the AGC-SF (left), each stage consists of a temporal low pass filter with a defined time constant and a 3-point spectral smoothing filter. The internal structure of an AGC-SF (right), the input of the AGC-SF comes from the lower temporal filter with lower time constant as well as the accumulation of the IHC. The output goes to the next stage of the temporal filter. The spectral smoothing filter of one stage is connected to lateral channels.

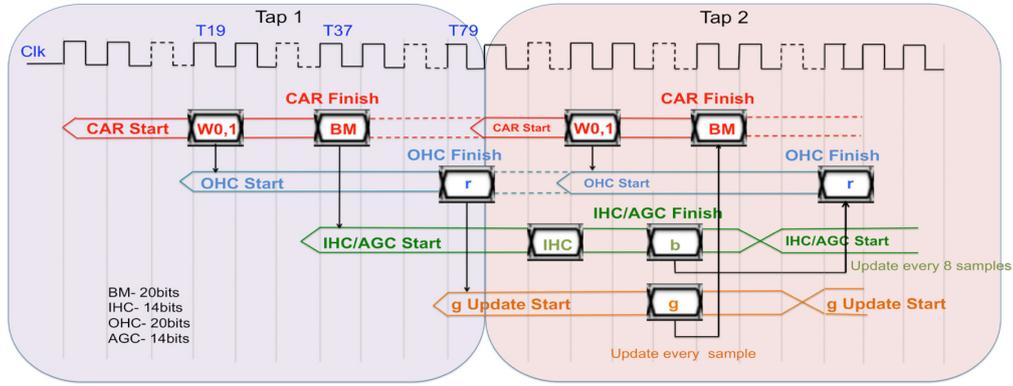


Fig. 5. Timing diagram of the CAR-FAC FPGA implementation.

The output of the IHC model is connected to the AGC loop-filter representing the spectrotemporal characterization of the medial olivo-cochlear (MOC) [19]. The AGC loop-filter is structured in parallel, as shown in Fig. 1, ensuring that the response of the AGC is fast, stable and non-ringing over a wide range of conditions. Furthermore, the AGC SFs operate at lower sampling rates than the CAR model as they function much slower than the accommodating sound frequencies at the CAR section. The AGC SFs also encompass lateral smoothing where the gains of the neighboring channels are maintained close to each other while their dynamic ranges are reduced.

### III. DESIGN METHODOLOGY AND FPGA IMPLEMENTATION

Each of the 100 channels comprises the BM segment as a part of the CAR model and the OHC, IHC and AGC sections as part of the FAC segment. The characteristic frequencies of the cascaded filters are set from 102 Hz to 12.673 kHz using the Greenwood mapping function [17]. Each BM filter is processed in a serial order of OHC, IHC, and AGC filters. The AGC filter output is fed back to the BM, thereby adapting its response to input sound levels.

#### A. Software Implementation

We initially implemented a floating-point simulation of the CAR-FAC model in Python. We then translated the design into a fixed-point implementation with a 20-bit word length size of the design parameters, 20-bit BM output, 20-bit OHC output, 14-bit IHC output, and 14-bit AGC.

#### B. FPGA Implementation

The CAR-FAC model is implemented on Altera Cyclone V FPGA. Its system clock frequency is 300 MHz and the input sound sampling rate is set at 32 kHz. This indicates that all the 100 sections are required to end their operation in less than 31.25  $\mu$ s. Our design pipelines the BM/OHC model and the IHC/AGC model to achieve real time operating speed. We use a time multiplexing approach to reuse hardware resources of one CAR-FAC section for all the 100 sections in the design. Fig. 5 displays the design timing-diagram of a single CAR-FAC section comprising the BM, OHC, IHC, and AGC filters respectively.

The model starts in the BM section and is followed by the OHC model when its internal states  $W_0$  and  $W_1$  are calculated and updated. The computation of the damping factor  $r$  in the OHC model signals the end of the current BM section processing and the start of the next section. Once  $BM_{out}$  is

obtained in the BM as section, it is passed into the IHC and AGC segment consisting of an HPF, an NLF and AGC smoothing filters (AGC-SF).

The AGC segment includes four LPFs, which perform computation at lower sampling rates. The number of LPFs is processed according to the down-sampling factor. Hence, at the 8<sup>th</sup> sample, AGC-SF1 invokes a single LPF; at every 16<sup>th</sup> sample, AGC-SF1 and AGC-SF2 invoke two LPFs; at every 64<sup>th</sup> sample, all four AGC-SFs invoke four LPFs. The output variable of the AGC,  $b$ , together with  $W_0$  and  $W_1$ , affects the variable  $r$  in the OHC segment. This varies the overall gain,  $g$ , as well as  $W_0$  and  $W_1$ , thereby resulting in the BM travelling wave amplitude conditioning.

The spatial filter in each section of the AGC is updated from the values of adjacent sections. For the first and last section filters, the input and output are respectively cycled back to the filter in place of non-existing neighbors. The NLF of IHC from (2) and (3) are implemented on FPGA as:

$$NLF_{IHC} = (1 - (BM\_hpf + 0.13)/4)^8 \quad (4)$$

The NLF of the OHC described in (1) is replaced by (5) as follows for hardware implementation, as the two equations are nearly identical over the expected range of values for  $v$ :

$$NLF_{OHC} = (1 - ((0.1v + 0.04)^2)/8)^8 \quad (5)$$

### IV. RESULTS

The FAC segment in the CAR-FAC model can be disabled in our design, thereby allowing flexibility in utilizing the FAC part. The CAR-FAC model is simulated in Cadence NCSim. Fig. 5 shows a comparison of the frequency response to a sine tone sweep from 20 Hz to 16 kHz at 32 kHz sampling rate for a 100-channel CAR and a 100-channel CAR-FAC model, respectively. For the latter, the model was first adapted to a 1 kHz sine wave of amplitude 0.05 and duration of one second, after which the AGC feedback parameter,  $b$ , was frozen. Fig. 7 illustrates the IHC output of the 100-channel CAR-FAC model for the same case. The reduction in gain around the adapted frequency is clearly visible in both figures. Fig. 8 shows the frozen AGC parameter,  $b$ , across 100 sections.

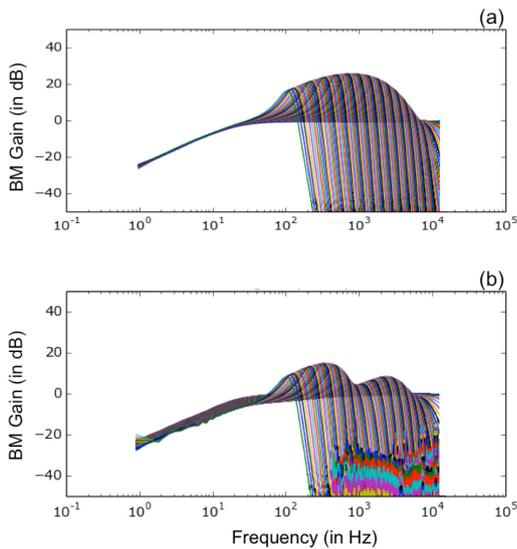


Fig. 6. Hardware implementation. BM frequency response for (a) CAR model; (b) CAR-FAC model.

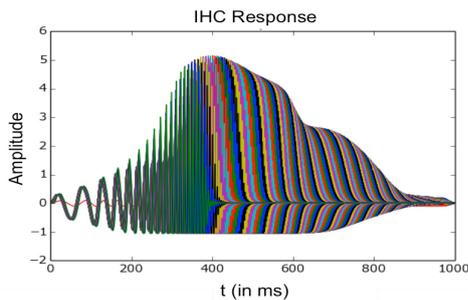


Fig. 7. Hardware implementation. 100-channel IHC output for the CAR-FAC model.

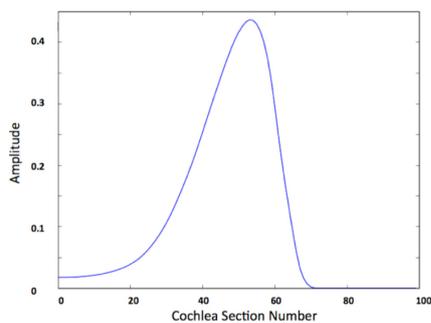


Fig. 8. Effect on parameter 'b' of the AGC in response to a 1 kHz pure tone across 100 sections.

## V. CONCLUSIONS

We have presented the design for an FPGA implementation of the fast acting compression (FAC) segment of the CAR-FAC model in this paper as an extension to our earlier work on the CAR model. We achieve this using the Altera Cyclone V FPGA running 100 frequency channels at 32 kHz sound sampling rate. The implementation of the CAR-FAC model on FPGA offers a real-time adaptive cochlear model and provides an interface for incorporating auditory models with on-the-fly speech and music processing [20], as well as for cross-sensory perception processing.

## REFERENCES

- [1] R. Meddis and E. A. Lopez-poveda, "Overview," in *Computational Models of the Auditory System*, R. Meddis, E. A. Lopez-Poveda, R. R. Fay, and A. N. Popper, Eds. Springer, 2010, pp. 1–6.
- [2] R. F. Lyon, A. G. Katsiamis, M. Park, A. Ox, and E. M. Drakakis, "History and Future of Auditory Filter Models," pp. 3809–3812, 2010.
- [3] R. F. Lyon, "Cascades of two-pole-two-zero asymmetric resonators are good models of peripheral auditory function.," *J. Acoust. Soc. Am.*, vol. 130, no. 6, pp. 3893–904, Dec. 2011.
- [4] J. O. Pickles, "The Outer and Middle Ears," in *An Introduction to the Physiology of Hearing*, 3rd ed., Brisbane, Queensland: Emerald Group, 2008, pp. 11–24.
- [5] B. M. Johnstone, R. Patuzzi, and G. K. Yates, "Basilar membrane measurements and the travelling wave," *Hear. Res.*, vol. 22, no. 1–3, pp. 147–153, 1986.
- [6] J. O. Pickles, "The Auditory Nerve," in *An Introduction to the Physiology of Hearing*, 3rd ed., Brisbane, Queensland: Emerald Group, 2008, pp. 73–101.
- [7] J. O. Pickles, "Mechanisms of Transduction and Excitation in the Cochlea," in *An Introduction to the Physiology of Hearing*, 3rd ed., Brisbane, Queensland: Emerald Group, 2008, pp. 103–152.
- [8] A. N. Lukashkin, M. N. Walling, and I. J. Russell, "Power Amplification in the Mammalian Cochlea," *Curr. Biol.*, vol. 17, no. 15, pp. 1340–1344, 2007.
- [9] R. K. Singh, "RTAP: Towards a Real-Time Auditory Periphery Simulation," *Int. Conf. Futur. Comput. Technol.*, pp. 52 – 57, 2015.
- [10] S. C. Liu, A. Van Schaik, B. A. Minch, and T. Delbruck, "Event-based 64-channel binaural silicon cochlea with Q enhancement mechanisms," in *ISCAS 2010 - 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems*, 2010, pp. 2027–2030.
- [11] V. Chan, S. C. Liu, and A. van Schaik, "AER EAR: A matched silicon cochlea pair with address event representation interface," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 54, no. 1, pp. 48–59, 2007.
- [12] T. J. Hamilton, C. Jin, A. van Schaik, and J. Tapson, "An active 2-D silicon cochlea," *IEEE Trans. Biomed. Circuits Syst.*, vol. 2, no. 1, pp. 30–43, 2008.
- [13] S. C. Liu, A. Van Schaik, B. A. Minch, and T. Delbruck, "Asynchronous binaural spatial audition sensor with 2X64X4 Channel output," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 4, pp. 453–464, 2014.
- [14] A. van Schaik, E. Fragnière, and E. A. Vittoz, "Improved Silicon Cochlea using Compatible Lateral Bipolar Transistors," *Adv. Neural Inf. Process. Syst.*, pp. 671–677, 1996.
- [15] A. van Schaik, T. J. Hamilton, and C. Jin, "Silicon models of the auditory pathway," in *Springer handbook of auditory research: computational models of the auditory system*, vol. 35, R. Meddis, E. A. Lopez-Poveda, R. R. Fay, and A. N. Popper, Eds. Springer, 2010.
- [16] R. M. Wang, T. J. Hamilton, J. C. Tapson, and A. van Schaik, "A neuromorphic implementation of multiple spike-timing synaptic plasticity rules for large-scale neural networks," *Front. Neurosci.*, vol. 9, no. May, pp. 1–17, 2015.
- [17] C. S. Thakur, T. J. Hamilton, J. Tapson, A. van Schaik, and R. F. Lyon, "FPGA implementation of the CAR Model of the cochlea," in *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014, pp. 1853–1856.
- [18] C. S. Thakur, T. J. Hamilton, J. Tapson, A. van Schaik, and R. F. Lyon, "Live Demonstration: FPGA implementation of the CAR Model of the cochlea," in *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014, pp. 1853–1856.
- [19] J. J. Guinan, "Olivocochlear Efferents: Anatomy, Physiology, Function, and the Measurement of Efferent Effects in Humans," *Ear Hear.*, vol. 27, no. 6, pp. 589–607, 2006.
- [20] C. S. Thakur, R. M. Wang, S. Afshar, T. J. Hamilton, J. C. Tapson, S. A. Shamma, and A. van Schaik, "Sound Stream Segregation: a neuromorphic approach to solve the 'cocktail party problem' in real time," *Front. Neurosci.*, vol. 9, no. September, pp. 1–10, 2015.