

エルゴード的順序論理回路を用いた蝸牛モデル： 哺乳類の非線形音声処理特性の再現と 効率的なFPGA実装

岸本, 唯 / KISHIMOTO, Yui

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学研究科編

(巻 / Volume)

65

(開始ページ / Start Page)

1

(終了ページ / End Page)

2

(発行年 / Year)

2024-03-24

(URL)

<https://doi.org/10.15002/00030695>

エルゴード的順序論理回路を用いた蝸牛モデル： 哺乳類の非線形音声処理特性の再現と効率的な FPGA 実装

A NOVEL INTEGRATED COCHLEAR MODEL BASED ON ERGODIC SEQUENTIAL LOGIC DYNAMICS:
REPRODUCTION OF NONLINEAR SOUNDPROCESSING AND EFFICIENT FPGA IMPLEMENTATION

岸本唯

Yui KISHIMOTO

指導教員 鳥飼弘幸

法政大学大学院理工学研究科電気電子工学専攻修士課程

In this paper, a novel hardware-efficient integrated cochlear model, whose nonlinear dynamics is described by ergodic sequential logics, is presented. It is shown that the presented cochlear model can reproduce pitch shift and combination tone generation of a mammalian cochlea. In addition, the presented model is implemented by a field-programmable gate array (FPGA) and its operations are verified by experiments. It is then shown that the presented model consumes much less hardware resources and much less power compared to a standard ordinary differential equation model of cochlea.

Key Words : Cochlear model, Ergodic sequential logic, Field programmable gate array (FPGA)

1. はじめに

哺乳類の耳は一般的に、外耳、中耳、内耳から構成されており、中でも内耳に位置する蝸牛は、主要な音響処理を行う渦巻き状の器官である。[1]. 図 1 は、リンパ液で満たされ、基底膜を含む哺乳類の蝸牛の展開図である。基底膜は、入力された音の周波数によって振動する位置が変わる周波数選択性を持つ。例えば、高い（低い）周波数の音が入力されると、その音に近い（遠い）基底膜が振動する。蝸牛を中心とした聴覚系では非線形フィルタ特性、多重トーン抑制、適応特性、第 1 及び第 2 ピッチシフト効果など多岐にわたる非線形音声処理が行われている [1]~[3]. 図 2(a) は哺乳類の蝸牛における非線形音声処理の特性を示したもので、グラフは、 f_1 と $f_2 > f_1$ から成る結合音入力に対する基底膜のある部分のフーリエ振幅スペクトラムを表している。入力周波数 f_1 、 f_2 において基底膜振動は強い振幅スペクトラムを有しているが、蝸牛の非線形音声処理の影響で基底膜は $n_1 f_1 \pm n_2 f_2$ ($n_1, n_2 \in \{1, 2, \dots\}$)で構成される周波数においても振動を有している。つまり f_1 、 f_2 から構成される結合音入力が入る蝸牛の非線形性により $n_1 f_1 \pm n_2 f_2$ ($n_1, n_2 \in \{1, 2, \dots\}$)で構成される周波数の基底膜振動を発生させている。このような蝸牛の非線形音声処理を結合音現象と呼ぶ。これらの蝸牛の持つ様々な非線形音声処理を再現する手法として様々なモデルが提案されてきた[4]~[6].

本研究では、エルゴード的順序論理回路を用いた統合蝸牛モデルを提案する。その結果、哺乳類の蝸牛の非線形

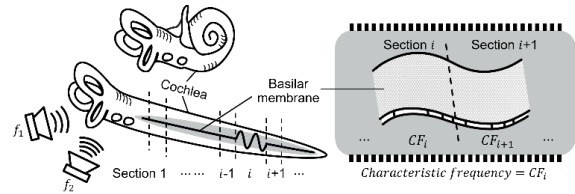


図 1 哺乳類の蝸牛の展開図。

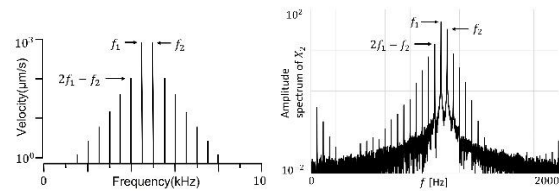


図 2 (a) 結合音現象の概要図 [2]. (b) 統合蝸牛モデルでのシミュレーションによる結合音現象の再現結果。

音声処理特性（結合音現象）を再現できることを示す（図 2(b)を参照）。さらに、このモデルは、標準的な ODE モデルに基づく統合蝸牛モデルと比較して、はるかにハードウェア効率がよいことを示す [7].

本研究の意義は非常に明確である：エルゴード的順序論理回路に基づく統合蝸牛モデルが、蝸牛の非線形音声処理特性を再現できる将来の人工内耳の設計に有用である。また、哺乳類の蝸牛をより厳密に模倣するには、より広い範囲の周波数領域に対応させる必要があるため、異なる特徴周波数をもつ複数の基底膜セクションモデルをつなげた統合モデルを設計することが重要である。

2. エルゴード的順序論理回路を用いた基底膜セクションモデル

(1) モデルの設計

本章では、エルゴード的順序論理回路を用いた基底膜セクションモデルを提案する。提案モデルは以下の4つの離散変数 $X_i, Y_i \in \mathbf{Z}_N = \{0, \dots, N-1\}$, $P_i, Q_i \in \mathbf{Z}_M = \{0, \dots, M-1\}$ を持つ。ここで、 i はモデルのインデックスである。 X_i と Y_i は離散状態変数で、非線形振動現象を起こす基底膜のふるまいを表している。 P_i と Q_i は離散補助変数で、状態依存分周期として動作する。 $N > 0$ かつ $M > 0$ は自然数で、提案モデルの解像度を決定するパラメータである。また、提案モデルは以下の2つの離散関数 $G_i, H_i \in \mathbf{Z}_N^2 \times \mathbf{Z}_M \rightarrow \{-1, 0, 1\}$ を持つ。関数 G_i および H_i は離散ベクトル場関数で、提案モデルの非線形ベクトル場を決定する。また、提案モデルは、 $t = 0, 1, 2, \dots$ のとき $C(t) = 1$, その他の場合 $C(t) = 0$ である内部クロック C と、2つの2値スイッチ信号 $S_X, S_Y \in \{0, 1\}$ によって動作する。また、3値の入力信号 $U \in \{-1, 0, 1\}$ を受け取る。これは提案モデルへの音刺激に対応している。内部クロック C は2値スイッチ信号 (S_X, S_Y) と離散ベクトル場関数 (G_i, H_i)、入力信号 U に依存した以下の離散状態変数 (X_i, Y_i) の遷移をもたらす。

If $C(t) = 1$, then

$$\begin{cases} X_i(t^+) = X_i(t) + S_X(t)G_i(X_i(t), Y_i(t), P_i(t)) + U_i, \\ Y_i(t^+) = Y_i(t) + S_Y(t)H_i(X_i(t), Y_i(t), Q_i(t)). \end{cases}$$

ここで、 $t^+ = \lim_{\varepsilon \rightarrow 0} t + \varepsilon$, $\varepsilon > 0$ である。これらの状態遷移において、離散状態変数 (X_i, Y_i) は最小値0, 最大値 $N-1$ において飽和する。

(2) 統合蝸牛モデルによる結合音現象の再現

次に、モデルに入力する信号を

$$U_1 = PDM(u), \quad U_2 = PDM(X_1)$$

とする。音入力 $u(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t)$ とし、 $A_1 > 0$, $A_2 > 0$ および $f_1 > 0$, $f_2 > 0$ はそれぞれ、入力信号 u の振幅、周波数である。音入力 u は3値入力信号 U_1 へ PDM を用いて密度変調され、1つ目のセクションモデルに入力される。 X_1 は1つ目の基底膜セクションモデルの出力である。 X_2 は2つ目の基底膜セクションモデルの出力を表し、これは2つのセクションを結合した統合蝸牛モデルの出力である。図2(b)は離散状態変数 X_2 のフーリエ振幅スペクトラムの数値シミュレーション波形である。図からもわかるように、入力周波数 f_1 および f_2 の他に、 $2f_1 - f_2$ をはじめとする $n_1 f_1 \pm n_2 f_2$ において強いピークを有しており、図2(a)のグラフと比較しても、結合音現象を再現しているといえる。

3. FPGA 実装とコスト比較

次に、提案モデルを FPGA へ実装する。離散状態変数 (X_i, Y_i) はビット長が $n = \lceil \log_2 N \rceil$ の符号付レジスタで実装され、離散補助変数 (P_i, Q_i) はビット長が $m = \lceil \log_2 M \rceil$ の符号付レジスタで実装される。さらに、離散ベクトル場関数 (G_i, H_i) はルックアップテーブルで実装される。また、比較のために、常微分方程式 (ODE) [8] の標準形に基づく統合蝸牛モデルも同じ FPGA ボードを使って実装した。その結果、提案するエルゴード的順序論理回路統合蝸牛モデルは、ODE 蝸牛モデルよりも少ない回路素子かつ低消費電力で実装できることが確認された。

4. まとめ

本研究では、エルゴード的順序論理回路に基づき、結合音現象を再現できるハードウェア効率のよい統合蝸牛モデルを提案した。提案モデルは、ODE モデルよりも少ない回路素子と消費電力で FPGA に実装可能であることを示した。この結果は、人間の内耳が持つ非線形の音声処理特性を知覚できる、小型かつ低消費電力の人工内耳開発に役立つと考えられる。今後の課題としては、蝸牛の他の非線形機能の再現や、より広い周波数範囲に対応する、より大きな統合蝸牛モデルの設計が挙げられる。

参考文献

- 1) J. O. Pickles, An Introduction to the Physiology of Hearing, 4th ed. Bingley, U.K.: Emerald Group, 2012.
- 2) M. A. Ruggero, "Responses to sound of the basilar membrane of the mammalian cochlea," Curr. Opin. Neurobiol., vol. 2, no. 4, pp. 449–456, 1992.
- 3) M. A. Ruggero, et al, "Basilar-membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am., vol. 101, no. 4, pp. 2151–2163, 1997.
- 4) R. Stoop, "From Hearing to Listening: Design and Properties of an Actively Tunable Electronic Hearing Sensor," Sensors, vol. 7, no. 12, pp. 3287–3298, 2007.
- 5) T. J. Hamilton et al., "A 2-D Cochlea with Hopf Oscillators," in Proc. The 2007 IEEE Biomedical Circuits and Systems Conference (BIOCAS), pp. 91–94, 2007.
- 6) K. Takeda and H. Torikai, "A Novel Hardware-Efficient cochlea model Based on Asynchronous Cellular Automaton Dynamics: Theoretical Analysis and FPGA Implementation," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 64, no. 9, pp. 1107–1111, 2017.
- 7) Y. Kishimoto and H. Torikai, "A hardware-efficient FPGA cochlear model for next generation nonlinear cochlear implant," in Proc. The 20th IEEE International SoC Design Conference (ISOCC), pp. 253–254, 2023.
- 8) A. Kern and R. Stoop, "Essential Role of Couplings between Hearing Nonlinearities," Phys. Rev. Lett., vol. 91, no. 12, 128101, 2003.