

Локализация источника звука искусственной нейронной сетью, основанной на модифицированных импульсных нейронах со следовой поляризацией

Р.А. Тикиджи – Хамбурьян, С.А. Полевая

Аннотация

В работе представлена искусственная нейронная сеть модифицированных импульсных нейронов (МИН) [10, 11], имитирующая локализацию источника звука при дихотической стимуляции. Модель содержит две группы МИН (“левую” и “правую”), которые получают возбуждающие и тормозные синапсы от соответствующих входных элементов, что делает каждый МИН аналогичным т.н. E-I нейронам в слуховой сенсорной системе [3, 4]. Показано, что разработанная модель может детектировать короткие (до 40 мкс.) временные задержки между сигналами при «дихотической» стимуляции и кодировать их как отношение количества спайков в ответах нейронных групп. Эффект ошибки при локализации источника звука, предсказанный при помощи имитационного моделирования и аналитического рассмотрения принципов детектирования, был на качественном уровне подтвержден экспериментально.

Sound-Source Localization by Neural Network Based on Modified Integrate-and –Fire Neuron Model with Autopolarization.

R.A. Tikidji – Hamburyan, S.A. Plevaya

Abstract

A neural-network model based on the modified "Integrate – and – Fire" neuron (MIFN) [10, 11] with autopolarization to simulate a sound localization at dichotic stimulation is presented. The model consists of two groups of MIFNs (“left” and “right”), which receive excitatory and inhibitory synapses from the corresponding input elements, making the each MIFN like to E-I neurons in the auditory sensory system [3, 4]. It has been shown that the model can detect a short (up to 40 μ s) Interaural Time Difference (ITD) by a ratio of spikes number in one neuron group to another. Model prediction about possible erroneous sound location received by computer simulation and analytical study was confirmed in psychophysical tests at qualitative level.

Введение

Хорошо известно, что слуховая система человека и животных способна различать короткие междушумные временные задержки (МВЗ) порядка 10 мкс, что позволяет ей локализовать источник звука в пространстве. Существуют несколько подходов к объяснению того, как нейронная система с характеристическими постоянными времени не менее 0.1 мс может детектировать столь короткие временные задержки. Доминирующей в настоящее время гипотезой, хорошо подтвержденной

экспериментально [1, 2, 3], является предположение о существовании детекторов совпадения (“coincidences detectors”), которые оперируют с частотно разделенными сигналами, имеющими фиксированную фазу [2, 7, 8, 9, 12]. В то же время имеются экспериментальные факты о наличии в слуховых отделах мозга трех специфических популяций нейронов, так называемых E-E, E-I и E-O нейронов [3, 5, 6]. На данный момент нет полного понимания вклада нейронов этих трех групп в детектирование МВЗ. Это касается в первую очередь E-I нейронов, которые наиболее широко представлены в слуховой коре (около 75% из всех зарегистрированных нейронов в первичной слуховой коре летучей мыши [5]). E-I клетки активируются при изолированной стимуляции контралатерального уха и, напротив, тормозятся при ипсилатеральной стимуляции. В данной работе представлена простая нейросетевая модель, имитирующая две популяции I-E нейронов и способная детектировать короткие МВЗ. Модель базируется на модифицированных импульсных нейронах (МИН) со следовой поляризацией (см. Приложение и <http://nisms.krinc.ru/Programmes/NeuroCADv201setup.exe>, [10, 11]). Показано, что в отличие от существующего представления об I-E нейронах как о грубом детекторе левой или правой стороны расположения источника звука, относительно большая популяция таких клеток способна детектировать МВЗ с высокой точностью (до 40 мкс), при условии, что любая постоянная времени у отдельных элементов сети больше, чем 1 мс.

1. Модель

Предлагаемая модель содержала две группы МИН по 100 клеток в каждой группе (Рис. 1). В предварительных численных экспериментах было обнаружено, что при меньшем количестве элементов точность детектирования резко снижается. Каждый МИН в группе имел параметры, случайным образом выбранные в пределах диапазона, указанного в Приложении, и получал два входа (синапса), от левого и правого входных (“сенсорных”) элементов, имитирующих билатеральные слуховые пути [1]. Клетки левой группы образовывали возбуждающий синапс с левым входным элементом и тормозный синапс - с правым. Для МИН правой популяции это правило было обратным. Вес (сила) каждого синапса может быть выбран по случайному закону в некотором диапазоне (см. Приложение). В данной работе для быстрого визуального анализа результатов вычислительных экспериментов распределение весов синапсов задавалось по линейному закону, при этом крайний левый элемент группы получал наиболее сильный тормозный вход и минимальное возбуждающее воздействие. Напротив, у крайнего правого элемента сила тормозного входа была наименьшей и он получал максимальное возбуждающее воздействие. Каждый входной элемент модели

являлся особой клеткой и был способен генерировать задаваемые последовательности импульсов. Выходом модели являлась импульсная активность всех нейронов обеих популяций и в виде развернутых во времени импульсных потоков, т.н. растрограмм (см. Рис 2(а),(б) и Рис 4(а) – (с)). Для интегральной оценки выхода модели определялось количество импульсов у всех нейронов в каждой из популяций при предъявлении единичного стимула.

2. Вычислительный эксперимент и аналитическое исследование

Для решения дифференциальных уравнений, описывающих МИН, был использован метод Рунге – Кутты второго порядка с временным шагом решения, равным половине минимальной МВЗ, что в большинстве случаев было эквивалентно 20 мкс. В компьютерных экспериментах имитировалась активность реальной слуховой системы при стимуляции короткими щелчками с широким спектром частот. При этом каждый входной стимул состоял из пары импульсов от левого и правого входных элементов, подаваемых с различной МВЗ (от -20мс до +20мс), анализировался ответ сети в течение 100 мс после предъявления стимула, затем модель инициализировалась заново.

Предварительные тесты показали, что правая и левая группы МИН в модели имеют разделенные минимумы активности (+1 мс и -1мс соответственно). Поэтому данный диапазон МВЗ был протестирован с более высоким временным разрешением. На рисунке 2а и 2б представлены ответы сети на стимулы с МВЗ, лежащей в диапазоне ± 1 мс. Видно, что увеличение или уменьшение МВЗ существенно влияет на активность “конкурирующих” групп нейронов. Для оценки этой зависимости определялось среднее количество импульсов в каждой группе при изменении МВЗ (рисунок 2в и 2г) и их отношение между конкурирующими группами. Так, для МВЗ равного -960мкс, это соотношение было равным 1.85, а для +400мкс соответственно 0.71. Аналогичные результаты были получены при активации сети с помощью воздействий, состоящих из пачек входных стимулов (100 Гц – 1кГц).

Для объяснения того, как данная модель “распознает” МВЗ рассмотрим элемент сети как простой импульсный интегрирующий нейрон, имеющий два входа разной полярности (Рис.3а). Потенциал интегратора будет описываться простым дифференциальным уравнением:

$$\tau \frac{du(t)}{dt} = \alpha(i_e - i_i) - \beta u(t), \quad (1)$$

где i_e и i_i возбуждающий и тормозный входной ток, $u(t)$ – мембранный потенциал, τ – постоянная времени интегратора, α и β – единичные коэффициенты, приводящие размерность. Для простоты рассмотрим случай, когда входные токи являются прямоугольными импульсами с длительностями τ_{se} и τ_{si} , имеют амплитуду I_e и I_i , соответственно, и разность фаз Δt . При условии, что $I_e > I_i$, $I_e - I_i > u(t)/\tau$, $\tau_{si} > \tau_{se}$ и возбуждающий импульс опережает тормозящий (т.е. $\Delta t < 0$), максимум $u(t)$ приходится на момент $t = \tau_{se}$; в обратном случае (т.е. $0 < \Delta t < \tau_{si} - \tau_{se}$), максимум приходится на момент времени $t = \Delta t + \tau_{se}$. Элементарным решением уравнения (1) при условии прямоугольных импульсов является:

$$u_{max} = \begin{cases} u(\tau_{se}) = I_e \left(1 - \exp\left(-\frac{\tau_{se}}{\tau}\right) \right) - I_i \left(1 - \exp\left(-\frac{(\tau_{se} - \Delta t)}{\tau}\right) \right) & \text{если } \Delta t < 0 \\ u(\tau_{se} + \Delta t) = I_e \left(1 - \exp\left(-\frac{\tau_{se}}{\tau}\right) \right) - I_i \left(1 - \exp\left(-\frac{(\tau_{se} + \Delta t)}{\tau}\right) \right) & \text{если } 0 < \Delta t < \tau_{si} - \tau_{se} \\ u(\tau_{se} + \Delta t) = I_e \left(1 - \exp\left(-\frac{\tau_{se}}{\tau}\right) \right) - I_i \left(1 - \exp\left(-\frac{\tau_{si}}{\tau}\right) \right) \exp\left(-\frac{(\tau_{si} - \Delta t - \tau_{se})}{\tau}\right) & \text{если } \tau_{si} - \tau_{se} < \Delta t \end{cases} \quad (2)$$

В соответствии с (2), максимальная амплитуда потенциала интегрирующего элемента зависит от длительности синаптического потенциала, постоянной времени интегрирующего элемента, веса возбуждающей и тормозной связей (т.е. синаптических токов) и задержки между входными сигналами (т.е. МВЗ). На рисунке 3б приведено семейство кривых $u_{max}(\Delta t)$. Легко видеть, что при коротких МВЗ лишь нейроны с $I_e \gg I_i$ могут достигнуть порога, однако при увеличении МВЗ все новые и новые нейроны популяции вовлекаются в импульсную активность. Таким образом, количество импульсов в нейронной группе может кодировать модуль МВЗ. Однако, $u_{max}(\Delta t)$ для нейронов конкурирующей группы имеет симметричный минимум, поскольку опережение и запаздывание возбуждающих и тормозных сигналов для этих групп противоположны, а следовательно $+\Delta t_r = -\Delta t_l$. Таким образом, минимумы $u_{max}(\Delta t)$, симметричные относительно нуля МВЗ, ограничивают диапазон точного кодирования коротких МВЗ, в котором отношение количества импульсов в конкурирующих группах кодирует МВЗ.

Рассмотрим случай, когда порог или мембранный потенциал у элементов групп смещены. Как следует из рисунка 3б, такие смещения приводят к изменению количества импульсов в правой и левой группах модели, что в свою очередь можно

представить как ошибочное кодирование МВЗ. Так, например, увеличение порога или уменьшение мембранного потенциала приводит к ошибке кодирования равносильной уменьшению МВЗ. Следовая гиперполяризация (СГП) и длительный тормозный постсинаптический потенциал (ДТПСП) аналогичным образом влияют на кодирование МВЗ, однако характер их влияния более сложный. В случае предъявления повторяющихся входных парных стимулов, влияние СГП и ДТПСП приводит к ошибочному кодированию, показанному на рисунках 4 а – в. Данный вариант ошибочного кодирования был получен при предъявлении серии из пяти пар стимулов, первая из которых имела задержку -900мкс, а последующие +900мкс. Временной интервал между стимулами в серии для результатов, представленных на рисунках 4а, 4б и 4в, составил 200мс, 600мс и 800мс соответственно. На рисунке 4г показано среднее количество импульсов в левой и правой группах МИН модели при различных интервалах между стимулами в сериях. Как следует из рисунка 4г, при уменьшении межстимульного интервала ошибка кодирования МВЗ возрастает. Этот феномен был верифицирован в психофизических тестах.

3. Результаты психофизического исследования

В исследовании участвовало 18 испытуемых в возрасте от 18 до 40 лет (девять мужчин и девять женщин). Испытуемых просили определить местоположения виртуального источника звука, образ которого создавался при помощи стимула состоящего из двух дипольных щелчков короткой длительности (не более 90мкс) с широким частотным спектром. Щелчки генерировались при помощи звуковой платы (Creative SB 128) под управлением компьютера на базе Intel Pentium III и подавались на левый и правый динамики головных стереонаушников (NAIKO CD3000). В идеальном случае такая система позволяет устанавливать разность фаз между щелчками в паре с точностью до 22,6 мкс, что является достаточным для качественных экспериментов подобного рода. Во время исследования задержка между щелчками в стимуле (т.е. МВЗ) была фиксирована и составляла +900мкс (опережение на левую сторону) или -900мкс (опережение на правую сторону). Повторяющиеся стимулы предъявлялись в виде серии, у которой межстимульные интервалы были постоянными. В каждой серии первые пять стимулов имели МВЗ равную -900мкс, а в последующих десяти парах МВЗ составляла +900мкс. Испытуемые указывали местонахождение каждого виртуального источника звука для каждой пары щелчков в азимутальных координатах. Усредненные результаты для серий с межстимульными интервалами 50мс, 100мс, 200мс, 300мс, 400мс, 500мс, 600мс, 700мс и 1с представлены на рисунке 4д. Как видно из приведенных результатов, ошибка в

восприятию виртуального источника звука возрастает с уменьшением межстимульного интервала, что согласуется с результатами, полученными в вычислительных экспериментах.

Заключение

Результаты вычислительных экспериментов и аналитическое исследование показали, что представленная модель, основанная на двух конкурирующих группах МИН, каждый из которых моделирует единичный E-I нейрон, способна детектировать короткие (до 40 мкс) МВЗ с высокой точностью. Точное кодирование МВЗ в модели возможно только в случае, если интервалы между парными стимулами в их серии порядка 1 с. При предъявлении пар стимулов с меньшим интервалом между ними наблюдалась ошибочная локализация звука и величина ошибки возрастает с уменьшением данного интервала. Феномен, обнаруженный в вычислительном эксперименте, был подтвержден в психофизическом исследовании на качественном уровне. Результаты моделирования и экспериментального исследования позволяют предположить, что ошибочная локализация источника звука в слуховой системе человека может быть обусловлена тормозными следовыми процессами на нейронном уровне (такими, как СГП и ДТПСР).

Представленная модель детектирует и кодирует МВЗ за счет интегрирования возбуждающего и тормозного воздействий единичным нейроном, подобно модели R. Izak с соавторами [4], и основана на оценке активности двух конкурирующих групп. Можно предположить, что данный механизм может вносить вклад в детектирование локализации источника звука, дополняя хорошо известную систему на основе детекторов совпадения [1, 2, 7 – 9, 12]. Вполне возможно, что подобные механизмы могут участвовать в процессах обработки коротких фазовых задержек и детекции рассогласования в других отделах нервной системы, например, в зрительной коре (в том числе, при детектировании нелинейности движения) и в коре мозжечка (при генерации сигналов коррекции ошибок в управлении движениями).

Авторы выражают глубокую благодарность Л.Н. Подладчиковой и А.И. Самарину за поддержку, помощь и дискуссию. Работа поддержана грантом РФФИ № 03-04-48369.

Приложение

Модифицированный импульсный нейрон описывается системой условных дифференциальных уравнений:

$$\left\{ \begin{array}{l}
 o(t) = 1, t' = t \quad \text{если } u(t) > \mathcal{G} \\
 \\
 o(t) = 1, a > 1, \\
 \frac{du(t)}{dt} = \frac{1}{C} I(t) + \frac{u_{ap}(t) - u(t)}{CR_{ap}} - \frac{u(t)}{\tau_s} + \frac{2U_s}{\tau_{fire}}, \\
 \frac{du_{ap}(t)}{dt} = \frac{u(t) - u_{ap}(t)}{\tau_{ap}} + \frac{U_{ap}}{\tau_{fire}}, \quad \text{если } t - t' < \frac{\tau_{fire}}{2} \\
 \\
 o(t) = 1, a > 1, \\
 \frac{du(t)}{dt} = \frac{1}{C} I(t) + \frac{u_{ap}(t) - u(t)}{CR_{ap}} - \frac{u(t)}{\tau_s} - \frac{2U_s}{\tau_{fire}}, \\
 \frac{du_{ap}(t)}{dt} = \frac{u(t) - u_{ap}(t)}{\tau_{ap}} + \frac{U_{ap}}{\tau_{fire}}, \quad \text{если } \frac{\tau_{fire}}{2} < t - t' < \tau_{fire} \\
 \\
 o(t) = 0, a = 1, \\
 \frac{du(t)}{dt} = \frac{1}{C} I(t) + \frac{u_{ap}(t) - u(t)}{CR_{ap}} - \frac{u(t)}{\tau}, \\
 \frac{du_{ap}(t)}{dt} = \frac{u(t) - u_{ap}(t)}{\tau_{ap}}, \quad \text{если } t - t' < \tau_{ref} \\
 \\
 o(t) = 0, a = 1, \\
 \frac{du(t)}{dt} = \frac{1}{C} I(t) + \frac{u_{ap}(t) - u(t)}{CR_{ap}} - \frac{u(t)}{\tau}, \quad \text{во всех остальных} \\
 \frac{du_{ap}(t)}{dt} = \frac{u(t) - u_{ap}(t)}{\tau_{ap}}, \quad \text{случаях}
 \end{array} \right.$$

$$\frac{du_i(t)}{dt} = \begin{cases} \frac{a}{\tau_r} (u(t) - u_i(t)) & \text{если } u(t) > u_i(t) \\ \frac{a}{\tau_f} (u(t) - u_i(t)) & \text{если } u(t) < u_i(t) \end{cases}$$

где: $o(t)$ – бинарная выходная функция нейрона, τ_{fire} и τ_{ref} – длительность импульса и период запрещения генерации импульсов¹ соответственно, $u(t)$ – мембранный потенциал (МП), C – емкость, $\tau = RC$ – характеристическое время интегратора, $I(t)$ – входной ток нейрона, R_{ap} – сопротивление связи, $\tau_{ap} = R_{ap} C_{ap}$ – характеристическое время дополнительного интегратора, U_{ap} – единичный следовый потенциал (СП), U_s – потенциал импульса, \mathcal{G} – порог импульсного разряда нейрона, a – коэффициент уменьшения постоянной времени во время импульса, c_{th} – постоянная составляющая порогового уровня МП, τ_r и τ_f – характеристические времена нарастания и спада порога, соответственно.

¹ период запрещения генерации импульсов - рефрактерный период

Связь нейронов описывается как динамический процесс выброса и истечения-разложения медиатора:

$$\frac{dm_i(t)}{dt} = \begin{cases} \frac{m_s - m_i(t)}{t_r - t_f} & \text{если } t - t'' < 2t_r \\ -\frac{m_i(t)}{t_f} & \text{если } t - t'' > 2t_r \end{cases}$$

где: m_i – количество медиатора в синапсе, m_s – количество медиатора при единичной инъекции, t'' – время от начала инъекции, t_r и t_f – характеристические постоянные времени выброса и стока медиатора.

Суммарный синаптический ток нейрона определяется как:

$$I(t) = \sum_N \omega_i m_i(t)$$

где: ω_i – весовой коэффициент данного синапса, N – количество синаптических контактов данного нейрона.

При построении модели параметры МИН выбирались из следующих диапазонов:

Параметр	Минимум	Максимум	Размерность
Характеристическое время основного интегратора	5	7.5	Мс
Емкость основного интегратора	15	20	пФ
Сопротивление связи	58	98	МОм
Емкость дополнительного интегратора	25	37	пФ
Амплитуда единичного СП	-10		мВ
Постоянная составляющая порога	6		
Характеристическая постоянная времени нарастания и спада порога	0.5	0.7	с.

Параметры связей выбирались из следующих диапазонов:

Параметр	Минимум	Максимум	Размерность
Для возбуждающих синапсов			
Синаптический вес (при условии $m_s = 1$)	290	770	пА
Характеристическая постоянная времени выброса медиатора	2		мс
Характеристическая постоянная времени стока – разложения медиатора	3	5	мс

Параметр	Минимум	Максимум	Размерность
Для тормозных синапсов			
Синаптический вес (при условии $m_s = 1$)	-540	-200	пА
Характеристическая постоянная времени выброса медиатора	3	4	мс
Характеристическая постоянная времени стока – разложения медиатора	30	50	мс

Список литературы

1. Gerstner W., Kistler W.M., Spiking Neuron Models. Single Neurons, Populations, Plasticity (Cambridge University Press, 2002)
2. Gold J.I., Knudsen E.I., A Site of Auditory Experience-Dependent Plasticity in the Neural Representation of Auditory Space in the Barn Owl s Inferior Colliculus, The Jour. of Neuroscience 20(9) (May 1, 2000) 3469-3486
3. Grothe B., Park T.J. , Sensitivity to Interaural Time Differences in the Medial Superior Olive of a Small Mammal, the Mexican Free-Tailed Bat, The Jour. of Neuroscience 18(16) (August 15, 1998) 6608-6622
4. Izak R., Scarbata G., Paschke P., Sound source localization with an Integrate-and-Fire neural system, in proc.: Seventh Inter. Conf. On Microelectronisc for Neural, Fuzzy and Bio-Inspired systems (Granada, Spain, April 1999) p.103
5. Liu W., Suga N., Binaural and commissural organization of the primary auditory cortex of the mustached bat, J. Comp. Physiol. A (1997) 181 599-605
6. Middlebrooks J.C., Xu L., Eddins A.C., Green D.M., Codes for Sound-Source Location in Nontonotopic Auditory Cortex, J. Neurophysiol. 80(1998) 863 - 881
7. Pen J.L. Viète S., Funabiki K., Saberi K., Konishi M., Cochlear and Neural Delays for Coincidence Detection in Owls, The Jour. of Neuroscience 21(23) (December 1, 2001) 9455-9459
8. Proctor L., Konishi M. , Representation of sound localization cues in the auditory thalamus of the barn owl, Neurobiology Vol. 94(September 1997) 10421-10425
9. Saberi K., Farahbod H., Konishi M. , How do owls localize interaurally phase-ambiguous

signals?, *Neurobiology* Vol. 95(May 1998) 6465-6468

10. Tikidji-Hamburyan, R.A. (2002) Thalamic circuitry model based on modified "Integrate-and-Fire Neurons", in proc.: of the 9th ICoNIP (Singapore, November 2002) pp. 576-581.
11. Tikidji-Hamburyan, R.A. Modified integrate-and-fire neuron as basic model for biologically plausible neural network. *Neurocomputers: design and application*, № 7-8, 2002, pp. 97-103.
12. Viete S., Pen J.L., Konishi M. , Effects of Interaural Intensity Difference on the Processing of Interaural Time Difference in the Owl s Nucleus Laminaris, *The Jour. of Neuroscience* 17(5) (March 1, 1997) 1815-1824