

На правах рукописи

КАЗАНОВИЧ  
Яков Борисович

**Осцилляторные нейросетевые модели  
когнитивных функций мозга**

Специальность 03.01.02 – Биофизика

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
доктора физико-математических наук

Пушино-2018

Работа выполнена в Лаборатории нейронных сетей Института математических проблем биологии РАН – филиале федерального государственного учреждения «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук»

**Научный консультант:** доктор физико-математических наук  
**Борисюк Роман Матвеевич**

**Официальные оппоненты:** доктор физико-математических наук  
**Редько Владимир Георгиевич**  
(гл.н.с., зам. рук. Центра оптико-нейронных технологий ФГУ ФНЦ НИИСИ РАН)

доктор физико-математических наук, профессор  
**Яхно Владимир Григорьевич**  
(зав. лаб. автоволновых процессов  
ФГБУН ФИЦ ИПФ РАН)

доктор биологических наук, профессор  
**Фролов Александр Алексеевич**  
(гл.н.с., зав. лаб. математической нейробиологии  
обучения ФГБУН ИВНД и НФ РАН)

**Ведущая организация:** Федеральное государственное автономное образовательное учреждение высшего образования «**Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского**», Институт биологии и биомедицины, кафедра нейротехнологий

Защита состоится «\_\_\_\_\_» \_\_\_\_\_ 2018 г. в \_\_\_\_\_ часов \_\_\_\_\_ минут на заседании совета Д 002.093.01 по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук при Федеральном государственном бюджетном учреждении науки Институте теоретической и экспериментальной биофизики Российской академии наук по адресу: ул. Институтская, 3, г. Пущино, 142290, Московская область.

С диссертацией можно ознакомиться в Центральной библиотеке ПНЦ РАН по адресу: ул. Институтская, 3, г. Пущино, 142290, Московская область, и на сайте ИТЭБ РАН: <http://web.iteb.psn.ru>.

Автореферат разослан «\_\_\_\_\_» \_\_\_\_\_ 2018 г.

Ученый секретарь  
диссертационного совета,  
кандидат физ.-мат. наук

Ланина Н.Ф.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Актуальность темы диссертации

В настоящее время нейробиология является научным направлением, где происходит быстрое накопление новых данных. Это дает надежду, что в недалеком будущем будет достигнут значительный прогресс в одной из наиболее важных проблем нейробиологии – понимании механизмов реализации основных когнитивных функций мозга человека, таких как сознание, мышление, внимание, память, ориентация в пространстве и др. Стимулом к этому служат желание усовершенствовать когнитивные способности человека, потребности, диктуемые практическими запросами медицины, и поиск новых путей создания “интеллектуальных” технических устройств.

Перспективным направлением для решения этих задач является исследование роли колебательных процессов в электрической активности структур мозга при выполнении различных когнитивных функций. Наличие ритмических компонент в различных частотных диапазонах хорошо известно со времен изобретения электроэнцефалографии (ЭЭГ). За прошедшие годы накоплен большой экспериментальный материал, убедительно показывающий, что тип колебаний и доминирующая в них частота коррелируют как с типом внешней стимуляции, так и с внутренним психологическим состоянием организма. Устойчивые пространственно-временные паттерны активности были обнаружены в различных областях мозга в экспериментах *in vivo* и *in vitro* на уровне одиночных нейронов, нейронных ансамблей и структур (областей) мозга. Эти структуры включают первичные зоны зрительной коры, обонятельную кору, сенсорно-моторную кору, таламус, гиппокамп, мозжечок и др. области [Basar and Bullock (Eds) // *Induced rhythms in the brain*, 1989; Steriade et al. // *Thalamic oscillations and signalling* 1990; Basar and Bullock (Eds) // *Brain dynamics. Progress and perspectives*. 1992; Gray // *Comput. Neurosci.*, 1994; Vinogradova // *Progr. Neurobiol.*, 1995; Basar // *Brain function and oscillations*, 1998; Freeman // *How brains make up their minds*, 2001; Buzsaki and Draguhn // *Science*, 2004]

Не удивительно, что колебательная активность имеет место при нейронном управлении периодическими движениями: дыханием, сердцебиением, перемещением в пространстве (ходьбе, плавании), однако необходимость колебаний в реализации когнитивных функций мозга не столь очевидна. Хотя все больше исследователей соглашаются с тем, что ритмическая активность и различные виды ее синхронизации являются универсальным механизмом координации работы мозга при кодировании, передаче и обработке информации [Pareti and De Palma // *Neurol. Sci.*, 2004], динамическая теория мозга все еще находится на начальном этапе своего развития, а многие возникшие в ней вопросы остаются без ответа.

Список функций мозга, в которых предположительно ведущую роль играют электродинамические процессы, включает:

- пространственную организацию обработки информации,

- связывание признаков (интеграция признаков объекта в цельный образ),
- селективное внимание,
- память (кратковременная и долговременная память, вспоминание),
- детекцию новизны,
- распознавание образов,
- адаптивное поведение,
- доминанту Ухтомского,
- сознание.

Компьютерное моделирование является эффективным инструментом для проверки справедливости различных гипотез относительно роли динамики нейронной активности в обработке информации мозгом. С помощью математического и вычислительного аппарата можно получить ответы на следующие вопросы:

- Каковы условия возникновения колебательной активности и ее синхронизации?
- Какие типы обработки информации могут быть реализованы на основе колебаний и их синхронизации в нейроподобных структурах?
- Какие преимущества дает использование колебательной активности по сравнению с методами, основанными на усредненной активности нейронных структур?
- Как объяснить имеющиеся экспериментальные данные относительно динамики активности структур мозга? Какова роль различных типов динамики в реализации когнитивных функций?

Этим вопросам посвящен специальный раздел теории нейронных сетей - осцилляторные нейронные сети (ОНС), основанный на теории нелинейных динамических систем, примененной к сетям из нейроподобных осцилляторных элементов. С обзорными работами в этой области можно ознакомиться по отечественным [Борисюк и др. // Математическое моделирование, 1992; Абарбанель и др. // Успехи физических наук, 1996; Борисюк и др. // Успехи физических наук, 2002; Рабинович и Мюезинолу // Успехи физических наук, 2010] и зарубежным [Levine et al. // Oscillations in neural systems, 1999; Miller // Time and the brain, 2000; Wang // IEEE Trans. Neural Networks, 2005; Rabinovich et al. // Trends Cogn. Sci., 2015] публикациям.

Несмотря на то что имеются многочисленные прямые и косвенные экспериментальные данные о том, что колебательная активность и ее синхронизация играют центральную роль в реализации когнитивных функций мозга, осцилляторные модели для многих экспериментальных условий либо крайне немногочисленны, либо вообще отсутствуют. Это, в частности, относится к моделям сегментации информации, селективного внимания, неоднозначного восприятия, детекции новизны и памяти. Разработке осцилляторных версий этих моделей и соответствующего математического аппарата для исследования динамики ОНС различной архитектуры посвящена настоящая диссертация.

## Цели и задачи

Настоящая работа представляет собой математическое и компьютерное исследование роли синхронизации нейронной активности в когнитивных функциях мозга. Цель исследования - доказательство с помощью математического моделирования того факта, что синхронизация является универсальным механизмом, позволяющим мозгу эффективно обрабатывать информацию при решении разнообразных когнитивных задач.

Для реализации поставленной цели были сформулированы следующие задачи:

1. Разработать теорию сетей фазовых осцилляторов с центральным элементом как эффективный и перспективный математический аппарат для исследования динамических режимов, связанных с когнитивными функциями мозга.
  - Провести анализ и классификацию режимов синхронизации фазовых осцилляторов и условий их реализации.
  - Провести анализ устойчивости стационарных режимов синхронизации при различных архитектурах сетей фазовых осцилляторов.
  - Разработать осцилляторную модель, реализующую принцип "победитель-получает-все", в терминах обобщенных фазовых осцилляторов.
2. Разработать комплекс осцилляторных моделей когнитивных функций мозга, основанных на принципах синхронизации нейронной активности.
  - Модель внимания и исследовать режимы ее функционирования при различных условиях фокусировки внимания.
  - Модель детекции новизны в гиппокампе, позволяющую воспроизводить тоническую и фазическую реакции гиппокампа на новые и известные стимулы. Исследовать надежность функционирования модели в зависимости от заполнения памяти.
  - Модель последовательного выбора объектов на изображении, отражающую основные этапы обработки информации в зрительной системе.
  - Модель обработки цветных изображений, позволяющую компенсировать дефекты определения контуров объектов.
  - Модель зрительного поиска, позволяющую воспроизводить результаты психофизических экспериментов в задачах разной степени сложности.
  - Модель одновременного слежения за движением нескольких зрительных объектов, позволяющую воспроизвести результаты психофизических экспериментов по оценке числа ошибок идентификации целевых объектов среди объектов-дистракторов.
  - Модель восприятия неоднозначных изображений, позволяющую воспроизвести функцию распределения времен восприятия неккеровского куба.
  - Модель запоминания и вспоминания временных последовательностей, устраняющую неопределенность продолжения вспоминания за счет контекста.

## Научная новизна работы

ОНС являются объектом многочисленных исследований как с точки зрения теоретического описания режимов функционирования, так и для моделирования различных природных явлений, включая работу мозга. Тем не менее, многие аспекты теории и практического использования ОНС остаются без должного внимания. Так, до недавнего времени в теории фазовых осцилляторов основное внимание уделялось сетям с однородными (глобальными или локальными) связями. Наши работы в области теории сетей фазовых осцилляторов с центральным элементом были пионерскими. Приложения сетей фазовых осцилляторов к моделированию когнитивных функций мозга также весьма немногочисленны. Наш приоритет состоит в том, что на многих примерах мы продемонстрировали, что сети фазовых осцилляторов являются эффективным инструментом для построения моделей когнитивных функций мозга. Их преимущество состоит в возможности описывать сложные процессы обработки информации в мозге достаточно простыми моделями, решать ряд вопросов относительно режимов функционирования моделей аналитически и производить масштабные вычислительные эксперименты в разумное время на персональных компьютерах.

При разработке теории сетей фазовых осцилляторов с центральным элементом нами впервые проведен анализ и классификация режимов синхронизации и получены условия устойчивости режимов синхронизации. Нами впервые предложено обобщение фазового осциллятора с включением в набор переменных собственной частоты и амплитуды осциллятора. Это позволило впервые реализовать в сетях фазовых осцилляторов один из универсальных принципов обработки информации в мозге и в искусственных нейронных сетях – "победитель-получает-все".

Научную новизну нашей работы характеризуют следующие факты.

1. Мы впервые продемонстрировали на осцилляторной нейросетевой модели, что формирование и переключение фокуса внимания может описываться в терминах синхронизации между информационными и управляющими областями коры. Отметим, что это было сделано за несколько лет до того, как были получены прямые экспериментальные данные о связи внимания с синхронизацией нейронной активности.

2. Мы впервые показали, что детекция новизны в гиппокампе может быть связана с разреженным кодированием и адаптацией собственных частот гиппокампальных осцилляторов, работающих на частоте тета-ритма.

3. На основе ОНС фазовых осцилляторов мы впервые построили полномасштабную модель обработки изображений, объединяющую на единых осцилляторных принципах такие когнитивные функции, как интеграция признаков объектов, сегментация объектов на цветном изображении, внимание и детекция новизны.

4. Мы впервые показали, что результаты психофизических экспериментов по зрительному поиску могут быть объяснены с помощью модели синхронизации в сети обобщенных фазовых осцилляторов.

5. Мы впервые показали, что ошибки при слежении за движением целевых объектов среди объектов-дистракторов могут быть объяснены как следствие ограниченности фазового пространства, в котором происходит оперирование системы внимания.

6. Мы впервые показали, что гамма-распределение времен восприятия неккеровского куба может быть объяснено в терминах частичной синхронизации и ее бистабильности в системе фазовых осцилляторов с центральным элементом.

7. Мы впервые показали с помощью осцилляторной нейросетевой модели, что число ошибок при вспоминании запомненной последовательности можно снизить при учете контекста, в котором происходило запоминание.

### **Научная и практическая значимость работы**

Данная работа посвящена анализу информационных процессов в нервной системе и носит характер фундаментального исследования. Разработанный математический аппарат для исследования сетей фазовых осцилляторов имеет более широкую область применения: он может использоваться при разработке вычислительных устройств на основе оптических сетей и сетей на джозефсоновских контактах. Особое значение в этом отношении может иметь реализованный нами осцилляторный вариант процедуры "победитель-получает-все". Разработанные модели когнитивных функций мозга могут использоваться в качестве математического инструмента для исследований в области теоретической нейрофизиологии в вопросах, связанных с ролью колебательной активности и синхронизации в обработке информации мозгом. Модели внимания и комплексные модели анализа изображений могут быть использованы для построения систем искусственного интеллекта, в частности, систем компьютерного зрения.

### **Методология исследования**

Работа выполнена в рамках методологии нейроинформатики (вычислительной нейробиологии) и теории нейронных сетей. Она включает в себя построение математических моделей на основе имеющихся экспериментальных данных и их исследование методами математического анализа и компьютерных вычислений. Основные математические результаты работы получены для сетей фазовых осцилляторов с центральным элементом методами бифуркационного анализа и теории устойчивости. Большинство моделей выполнено на крупномасштабном уровне с использованием в качестве элементов моделей больших нейронных ансамблей, представленных фазовыми или обобщенными фазовыми осцилляторами. Исключение составляют модель запоминания временных последовательностей, построенная из нейронов типа Ходжкина-Хаксли, а также ряд моделей, описанных в Приложении. Модели в Приложении представляют собой аналоги некоторых моделей из основной части диссертации, но их реализация осуществлена с помощью ОНС из импульсных нейронов.

## **Основные положения (результаты), выносимые на защиту**

1. Разработка математического аппарата для исследования режимов синхронизации сетей фазовых осцилляторов с центральным элементом. Анализ и классификация режимов синхронизации и условий их устойчивости. Описание условий существования и устойчивости режима частичной синхронизации как одного из наиболее значимых видов нейронной активности при когнитивной деятельности.

2. Реализация процедуры "победитель-получает-все" в ОНС обобщенных фазовых осцилляторов с центральным элементом. Обобщение фазовых осцилляторов с включением в набор переменных собственных частот и амплитуд осцилляторов. Определение оптимальных параметров для реализации процедуры "победитель-получает-все".

3. Разработка осцилляторной модели внимания. Получение аналитических и численных результатов для обоснования гипотезы о том, что формирование и переключение фокуса внимания может описываться в терминах синхронизации между информационными и управляющими областями коры.

4. Разработка осцилляторной модели детекции новизны в гиппокампе. Демонстрация с помощью численных экспериментов, что детекция новизны в гиппокампе может быть реализована с использованием разреженного кодирования и адаптации собственных частот гиппокампальных осцилляторов, работающих на частоте тета-ритма.

5. Разработка осцилляторной модели последовательного выбора объектов на изображении, отражающей основные стадии обработки зрительной информации, начиная с первичных отделов зрительной коры и заканчивая детекцией новизны в гиппокампе. Демонстрация с помощью численных экспериментов того, что модель способна последовательно включать объекты, присутствующие в зрительном поле, в фокус внимания, выделять геометрические признаки объектов, инвариантные к положению и масштабу, и производить анализ объектов на новизну в терминах длительности (тонической или фазической) осцилляторного ответа гиппокампа.

6. Разработка осцилляторной модели сегментации цветных изображений. Демонстрация с помощью численных экспериментов того, что сочетание традиционных алгоритмов параллельной обработки изображений с методами, основанными на синхронизации, позволяет улучшить качество выделения контуров и сегментации отдельных объектов на изображении.

7. Разработка осцилляторной модели зрительного поиска. Демонстрация с помощью численных экспериментов того, что результаты психофизических экспериментов по зрительному поиску в задачах различной степени сложности могут быть объяснены с помощью модели синхронизации в сети обобщенных фазовых осцилляторов.

8. Разработка осцилляторной модели слежения за несколькими объектами. Получение аргументов, основанных на численных экспериментах, в пользу гипотезы, что ошибки при слежении за движением целевых объектов среди объектов-дистракторов могут быть объяснены как следствие



ограниченности фазового пространства, в котором происходит оперирование системы внимания.

9. Разработка осцилляторной модели восприятия неоднозначных изображений. Демонстрация с помощью численных экспериментов того, что гамма-распределение времен восприятия неккеровского куба может быть объяснено в терминах частичной синхронизации и ее бистабильности в системе фазовых осцилляторов с центральным элементом.

10. Разработка осцилляторной модели запоминания временных последовательностей. Демонстрация с помощью численных экспериментов того, что запоминание временных последовательностей может быть реализовано в ОНС с использованием синаптической модификации по принципам STDP и анти-STDP.

### **Апробация работы**

Результаты, полученные в диссертации, представлялись в виде докладов на отечественных и международных конференциях (опубликовано 2 приглашенных доклада, 20 статей докладов и 24 кратких тезиса докладов), а также на семинарах в отечественных и зарубежных научных учреждениях.

### **Публикации и вклад автора**

Результаты диссертации опубликованы в 30 рецензируемых журнальных статьях из списка ВАК (из них 21 в международных журналах). Имеется также 26 публикаций в рецензируемых журналах, не входящих в перечень ВАК, монографиях и сборниках научных статей (из них 16 в иностранных изданиях). Во всех совместных работах основные идеи, теоретический анализ и проведение основных вычислительных экспериментов принадлежат автору диссертации. Исключение составляют модели, представленные в Приложении, в которых вычислительные эксперименты были проведены соавторами.

### **Объем и структура диссертации.**

Диссертация состоит из 5 глав, Выводов и Приложения. Общий объем диссертации 463 страницы. Диссертация содержит 13 таблиц, 114 рисунков и список литературы из 485 наименований.

## **ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ**

Хотя исследования осцилляторной активности мозга имеют давнюю историю, привлечение к этой проблеме широкого внимания со стороны специалистов по математическому моделированию связано с экспериментами Зингера с сотрудниками [Gray et al. // Nature, 1989; Engel et al. // Proc. Natl. Acad. Sci. (USA), 1991] по проблеме интеграции признаков объектов в цельный образ. За минувшие годы был накоплен большой объем экспериментальных данных о связи когнитивных функций мозга с осцилляторной активностью в различных ритмических диапазонах и с синхронизацией этой активности. Для объяснения и воспроизведения этих феноменов возникла новая парадигма

нейросетевого моделирования – осцилляторные нейронные сети (ОНС). Основным математическим аппаратом ОНС – системы дифференциальных уравнений, описывающих динамику активности взаимодействующих нейронов или нейронных популяций. Такие модели могут исследоваться как аналитически методами теории бифуркаций, так и с помощью имитационных расчетов на компьютере.

Одной из характерных особенностей ОНС является способность к адаптации, проявляющаяся в модификации параметров элементов или связей в ответ на внешнюю стимуляцию. Эта модификация может привести к изменению динамического поведения сети по мере предъявления различных входных стимулов.

В случае, когда активность нейронного ансамбля имеет осцилляторный характер, удобно описывать эту активность посредством фазового осциллятора [Kuramoto // *Chemical oscillations, waves and turbulence*, 1984], динамика которого характеризуется лишь одной переменной – фазой колебания. Основная задача состоит в описании областей пространства параметров, соответствующих различным режимам синхронизации (полной или частичной). Такие модели широко используются для воспроизведения нейрофизиологических данных относительно синхронизации нейронной активности и функциональных связей в мозге [Sompolinsky et al. // *Proc. Natl. Acad. Sci.*, 1990; Schuster and Wagner // *Biol. Cybern.*, 1990; Kuramoto et al. // *Progr. Theor. Phys.*, 1992; Ermentrout and Kopell // *Neural Comput.*, 1994; Breakspear et al. // *Front. Hum. Neurosci.*, 2010; Gleiser and Spormaker // *Phil. Trans. R. Soc. A*, 2010].

Анализ литературы позволил сформулировать ряд вопросов, которые послужили исходным стимулом проводимых исследований, и предложить методологию, в рамках которой на эти вопросы могут быть получены ответы:

1. Колебательная активность и синхронизация нейронной активности сопровождают многие когнитивные функции и являются гипотетической основой реализации этих функций. Выяснение, каковы потенциальные возможности этих механизмов и как они согласуются с экспериментальными данными, является одной из основных целей моделирования когнитивных функций.

2. Имеется большое количество моделей с осцилляторной динамикой, реализующих те или иные когнитивные функции. Возникает вопрос о возможности универсального подхода к моделированию широкого спектра когнитивных функций. Для этого необходимо предложить соответствующий математический аппарат и продемонстрировать возможность его аналитического исследования и эффективного использования в вычислительных экспериментах.

3. В качестве такого аппарата предлагается использовать фазовые осцилляторы (обобщенные фазовые осцилляторы), которые уже успешно применялись при моделировании эффектов синхронизации в различных областях естественных наук. Фазовые осцилляторы полезны тем, что позволяют редуцировать большие ансамбли нейронов до простых

осцилляторных элементов, что значительно облегчает исследование сложных динамических взаимодействий в осцилляторных системах.

4. С точки зрения архитектуры нейронных сетей для реализации когнитивных функций особое внимание заслуживает архитектура с центральным управляющим элементом. Именно в такой архитектуре наиболее эффективно реализуется принцип "победитель-получает-все", являющийся универсальным для многих когнитивных функций. Архитектуры с центральным управляющим элементом до недавнего времени не были в фокусе внимания специалистов по моделированию, поэтому исследование их возможностей представляет собой новый фронт исследований.

5. Наблюдаемая во многих экспериментах модуляция амплитуды колебаний в зависимости от задач селекции входной информации естественным образом может быть реализована на принципах резонанса. Нейронный ансамбль, работающий синхронно с центральным управляющим элементом, увеличивает свою амплитуду (и усиливает взаимодействие с центральным управляющим элементом), в то время как активность других ансамблей (и возможность их влияния на центральный управляющий элемент) подавляется до низкого уровня.

6. Реализация памяти в нейронных сетях обычно осуществляется путем адаптации синаптических связей. Другим видом адаптации является изменение параметров самих элементов нейронной сети. Этот тип адаптации редко используется при моделировании. Мы покажем, что для ОНС адаптация собственной частоты осцилляторов является эффективным способом организации синхронного взаимодействия между центральным управляющим элементом и периферическими элементами сети.

Основное содержание диссертации сосредоточено в главах 3 и 4. Глава 3 посвящена разработке теории фазовых осцилляторов с центральным элементом. Результаты этой главы имеют самостоятельный интерес, а также используются при построении ряда моделей в главе 4. Глава 4 посвящена описанию разработанных моделей ОНС для реализации различных когнитивных функций и результатов соответствующих симуляционных вычислений.

### **Синхронизация в сети фазовых осцилляторов с центральным элементом**

Фазовые осцилляторы были предложены Курамото [Kuramoto // Lecture Notes in Physics, 1975; Kuramoto // Chemical oscillations, waves and turbulence, 1984] для описания синхронизации в колебательных системах. Они нашли применение в различных областях физики, химии и биологии.

Традиционно модель Курамото описывается следующими уравнениями:

$$\frac{d\theta_i}{dt} = \omega_i + \sum_{j=1}^n K_{ij} f(\theta_j - \theta_i), \quad i = 1, \dots, n, \quad (1)$$

где  $\theta_i$  - фаза  $i$ -го осциллятора,  $\omega_i$  - собственная частота осциллятора,  $f$  - функция взаимодействия между осцилляторами, нечетная, периодическая,  $K_{ij}$  - параметр связи, задающий силу воздействия  $j$ -го осциллятора на  $i$ -й

осциллятор. При отсутствии взаимодействия между осцилляторами они работают с частотой  $\omega_i$ .

Модель с синусоидальной функцией взаимодействия, в которую дополнительно входят фазовые сдвиги, имеет вид

$$\frac{d\theta_i}{dt} = \omega_i + \sum_{j=1}^n K_{ij} \sin(\theta_j - \theta_i - \alpha_{ij}), \quad i = 1, \dots, n$$

и называется моделью Курамото-Сакагучи.

Большинство результатов относительно синхронизации фазовых осцилляторов получено для сетей с глобальными или локальными связями. Системы с центральным элементом долгое время оставались вне поля зрения исследователей. Публикации на эту тему появились в основном в последние годы, будучи стимулированы нашими более ранними работами.

Мы опишем результаты по математической теории фазовых осцилляторов с центральным элементом, которые лежат в основе построения ряда моделей главы 4 диссертации.

Рассмотрим сеть фазовых осцилляторов Курамото-Сакагучи с центральным элементом. Все взаимодействия в этой системе осуществляются только через центральный осциллятор (ЦО). Нас будут интересовать два типа синхронизации, которые могут возникнуть в такой системе: 1) *полная синхронизация*, когда все осцилляторы работают с одной и той же частотой, и 2) *частичная синхронизация*, когда часть осцилляторов работает синхронно с ЦО, а остальные осцилляторы работают несинхронизованно. Тип синхронной активности определяется параметрами системы – величинами собственных частот осцилляторов и силами взаимодействия между осцилляторами. Заметим, что в отличие от традиционной модели Курамото рассматриваемая система обладает способностью перестраивать паттерн синхронизации в зависимости от собственной частоты центрального осциллятора.

Модель с ЦО описывается уравнениями

$$\frac{d\theta_0}{dt} = \omega_0 + \frac{A}{n} \sum_{j=1}^n f(\theta_j - \theta_0 + \gamma), \quad (2)$$

$$\frac{d\theta_i}{dt} = \omega_i + B g(\theta_0 - \theta_i + \delta), \quad i = 1, \dots, n, \quad (3)$$

где  $A$  и  $B$  – параметры связи (взаимодействия). Уравнение (2) описывает динамику ЦО. Осцилляторы, описываемые уравнением (3), называются периферическими осцилляторами (ПО). Без потери общности можно считать, что  $A > 0, B > 0, \delta = 0$ . Параметр  $\gamma$  может принимать как положительные (фазовое опережение), так и отрицательные (фазовое запаздывание) значения.

Вычитая (2) из (3) и полагая  $\varphi_i = \theta_i - \theta_0$  получаем уравнения в разностях фаз

$$\frac{d\varphi_i}{dt} = \omega_i - \omega_0 - Bg(\varphi_i) - \frac{A}{n} \sum_{j=1}^n f_i(\varphi_j). \quad (4)$$

Пусть  $f(x) = g(x) = \sin(x)$  и число ПО  $n$  велико (это в большинстве случаев соответствует нейробиологической реальности). В диссертации получены уравнения для определения частоты при полной или частичной синхронизации. Результаты решения этих уравнений при  $\omega_i$ , равномерно распределенных в интервале  $(-1, 1)$ , приведены на рис. 1, 2.

Стоит подчеркнуть, что полученные результаты дают потенциальную возможность для целенаправленного выбора параметров модели таким образом, чтобы обеспечить реализацию требуемого режима синхронизации, полной или частичной.

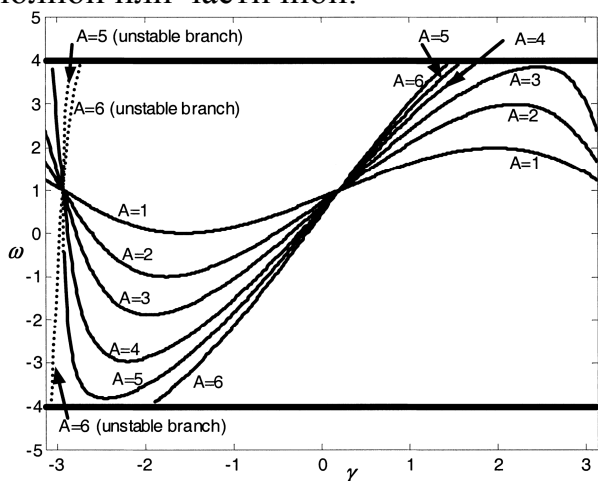


Рис. 1. Частота полной синхронизации как функция фазового сдвига для различных значений силы связи  $A$ . Сплошные и пунктирные линии соответствуют устойчивым и неустойчивым режимам полной синхронизации. Жирные горизонтальные линии ограничивают область существования полной синхронизации. Значения параметров:  $\omega_0 = 1, B = 5$ .

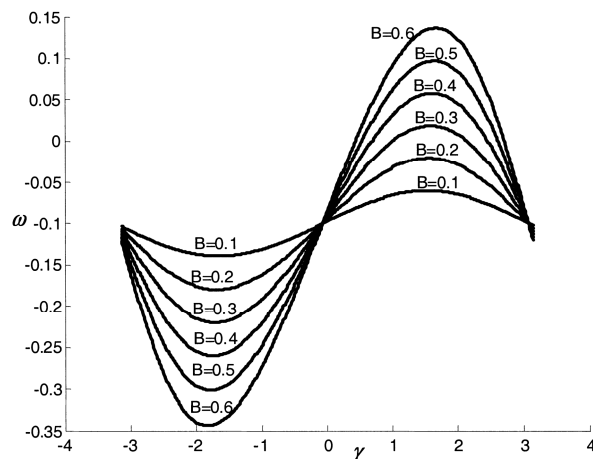


Рис. 2. Частота частичной синхронизации как функция фазового сдвига при различных значениях параметра взаимодействия  $B$ . Значения параметров системы:  $\omega_0 = -0,1, A = 0,5$ .

### Анализ устойчивости режимов синхронизации в сетях фазовых осцилляторов с центральным элементом

В отличие от предыдущего раздела, где в качестве функций взаимодействия использовалась функция  $\sin(x)$ , в данном разделе будет рассматриваться система идентичных фазовых осцилляторов  $\omega_0 = \omega_i = \omega$  с более общими функциями взаимодействия:

$$f'(0) = a_1, f'(\pi) = a_2, g'(0) = b_1, g'(\pi) = b_2, A = 1/n, B = 1. \quad (5)$$

Точки  $\Phi = (\varphi_1, \dots, \varphi_n)$  с координатами  $\varphi_i \in \{0, \pi\} (i = 1, \dots, n)$  являются стационарными точками системы (4). В диссертации получены условия устойчивости точек  $\Phi_k$ , у которых  $k$  координат равны 0 и  $n - k$  координат равны  $\pi$ .

На рис. 3 даны примеры областей устойчивости (неустойчивости) точек  $\Phi_k$  для различных значений параметров. В случае функций взаимодействия вида

$$f(x) = a(\sin(x) + r \sin(2x)), \quad g(x) = b(\sin(x) + p \sin(2x))$$

в диссертации приведены явные условия на параметры системы, обеспечивающие устойчивость точек  $\Phi_k$ .

При введении в систему (2)-(3), (5) локальных связей показано, что если  $n$  достаточно велико, даже небольшое десинхронизирующее воздействие от ЦО к ПО разрушает устойчивость точки  $\Phi_n = (0, \dots, 0)$  (полная синхронизация), однако наличие дальних связей позволяет сохранить устойчивость  $\Phi_n$ .

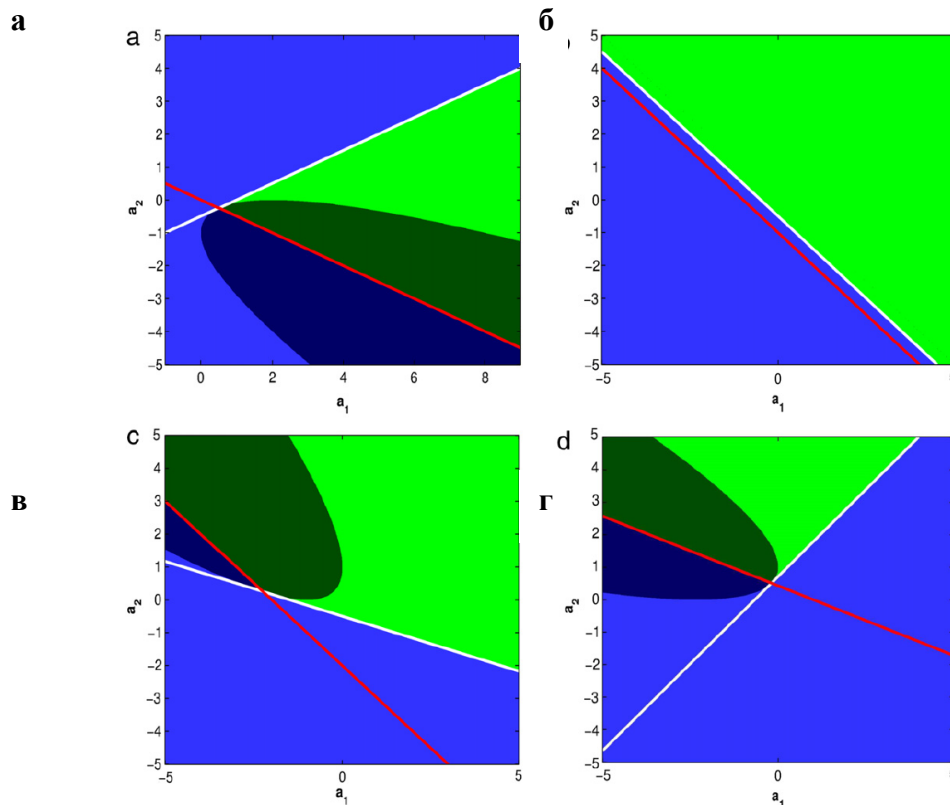


Рис. 3. Области устойчивости (неустойчивости) точки  $\Phi_k$ . Цвет показывает: зеленый – устойчивые узлы, темно-зеленый – устойчивые фокусы, синий – неустойчивые узлы, темно-синий – неустойчивые фокусы. Значения параметров: (а)  $n = 3, k = 1, b_1 = -1, b_2 = 1$ ; (б)  $n = 4, k = 2, b_1 = b_2 = 1$ ; (в)  $n = 4, k = 2, b_1 = 3, b_2 = 1$ ; (г)  $n = 10, k = 3, b_1 = 2, b_2 = 5$ .

Интерес к системе (2)-(3) и ее динамике связан с тем, что эта система была нами использована в качестве одного из вариантов модели внимания. Полученные результаты дают возможность ориентироваться в различных режимах синхронизации и соответствующих им режимах фокусировки внимания в зависимости от архитектуры системы и ее параметров.

### Процедура "победитель-получает-все" в сети обобщенных фазовых осцилляторов

Непосредственное использование системы (2)-(3), (5) как модели внимания ограничено тем обстоятельством, что помимо стационарных режимов значительную область пространства параметров занимают многочастотные и хаотические колебания, причем эти колебания сосуществуют со

стационарными состояниями различных режимов синхронизации. Для модели внимания требуется такое устойчивое функционирование системы, при котором можно гарантировать, что конкуренцию за синхронизацию с ЦО выиграет не более чем один ПО, а случаи, когда победителя не оказывается (что соответствует несформированному фокусу внимания), возникают, по возможности, редко. Для этого сеть с фиксированными параметрами не подходит. Мы покажем, что режим "победитель-получает-все" (ППВ) обеспечивается в сети обобщенных фазовых осцилляторов с центральным элементом и адаптирующимися параметрами. В отличие от системы Курамото, в которой осцилляторы описываются одной переменной (фазой колебаний), а собственные частоты и величины связей являются параметрами, в случае обобщенных осцилляторов эти параметры превращаются в динамические переменные.

ППВ является вычислительным принципом, широко используемым в искусственных нейронных сетях для реализации таких функций, как конкурентное обучение, принятие решений и выбор действия [Хайкин // Нейронные сети. Полный курс, 2006]. Традиционно системы, реализующие процедуру ППВ, строятся из искусственных пороговых нейронов или импульсных нейронов, которые конкурируют за уровень активности с помощью латерального или рекуррентного торможения. Предложенный нами альтернативный подход основан на синхронизации активности в сети обобщенных фазовых осцилляторов с радиальной архитектурой связей. ПО, выигравший конкуренцию за синхронизацию с ЦО, резонансно наращивает свою активность до высокого уровня, в то время как активность "проигравших" ПО подавляется до низкого уровня.

Динамика модели описывается уравнениями

$$\frac{d\theta_0}{dt} = \omega_0 + \frac{1}{n} \sum_{i=1}^n a_i f(\theta_i - \theta_0), \quad (6)$$

$$\frac{d\theta_i}{dt} = \omega_i + b g(\theta_0 - \theta_i), \quad i = 1, \dots, n, \quad (7)$$

$$\frac{d\omega_0}{dt} = \frac{\alpha}{n} \sum_{i=1}^n a_i f(\theta_i - \theta_0), \quad (8)$$

$$\frac{da_i}{dt} = \beta(-a_i + c + \gamma h(\theta_i - \theta_0)), \quad i = 1, \dots, n, \quad (9)$$

где  $\theta = (\theta_0, \theta_1, \dots, \theta_n) \in T^{n+1}$ ,  $\omega_0 \in R$ ,  $a = (a_1, \dots, a_n) \in R^n$  - переменные,  $\alpha, \beta, \gamma, c \in R$ ,  $b < 0$  - параметры. Без ограничения общности можно полагать  $b = -1$ . Таким образом, ПО оказывают на ЦО синхронизирующее воздействие, а ЦО десинхронизирует ПО. Связи от ПО к ЦО  $a_i$  будем ассоциировать с амплитудой колебаний ПО.

Функции  $f$ ,  $g$ ,  $h$   $2\pi$ -периодические и удовлетворяющие следующим условиям:

$$f(x) = -f(-x), f'(0) > 0, f'(\pi) = 0, g(x) = -g(-x), g'(0) > 0, g'(\pi) < 0,$$

$$h(x) = h(-x), h(0) = 1, h(\pi) = 0, h'(0) = h'(\pi) = 0 \quad (10)$$

Таким образом, функции  $f$ ,  $g$  – нечетные,  $h$  – четная. Из периодичности и нечетности  $f$  и  $g$  следует, что

$$f(0) = f(\pi) = g(0) = g(\pi) = 0.$$

Будем предполагать, что других нулей у этих функций нет. Будем также считать, что функция  $h(x)$  монотонно возрастает на  $[-\pi, 0]$  (и, соответственно, монотонно убывает на  $[0, \pi]$ ).

Уравнения (6)-(9) можно рассматривать как обобщение стандартных уравнений Курамото для фазовых осцилляторов, полученное путем введения дополнительно к фазовым уравнениям (6)-(7) уравнения (8) для адаптации собственной частоты ЦО и уравнения (9) для адаптации амплитуд ПО.

Смысл уравнения (8) становится ясным, если переписать его как

$$\frac{d\omega_0}{dt} = \alpha \left( \frac{d\theta_0}{dt} - \omega_0 \right).$$

В соответствии с этим уравнением собственная частота ЦО адаптируется по направлению к его текущей частоте. Параметр  $a$  задает скорость адаптации.

Смысл уравнения (9) состоит в резонансном возрастании амплитуды колебаний  $i$ -го ПО до уровня  $c + \gamma$ , если этот ПО работает синфазно с ЦО, в противном случае амплитуда ПО падает до низкого уровня  $c$ . Параметр  $\beta$  задает скорость изменения амплитуд.

В работе [Byrylko et al. // Sci. Rep., 2018] показано, что система (6)-(10) реализует процедуру ППВ, при этом не более чем один ПО может выиграть конкуренцию за синхронизацию с ЦО и увеличить свою амплитуду до резонансного уровня. В зависимости от соотношения между величиной  $b$  и разбросом собственных частот  $\omega_i$  режим ППВ может быть как стационарным, так и нестационарным.

При проведении вычислительных экспериментов в качестве функций  $f$ ,  $g$ ,  $h$  были использованы следующие функции:

$$f(x) = \begin{cases} \sin\left(\frac{|x - \pi|^v}{\pi^{v-1}}\right), & x \in [0, \pi], \\ -\sin\left(\frac{|x + \pi|^v}{\pi^{v-1}}\right), & x \in [-\pi, 0], \end{cases} \quad g(x) = \sin(x), \quad h(x) = \begin{cases} \left(\frac{\mu^2 - x^2}{\mu^2}\right)^\sigma, & |x| < \mu, \\ 0, & \mu \leq |x| \leq \pi \end{cases} \quad (11)$$

с параметрами  $v > 1, \sigma \gg 1, \mu \in (0, \pi)$  (рис. 4).

С помощью имитационного моделирования было показано, что оптимальные условия для функционирования модели в режиме ППВ реализуются при следующих условиях:

1. Достаточно высокие значения начальной амплитуды ПО, победа которого предпочтительна, по сравнению с начальными амплитудами других ПО.



2. Начальные значения фаз ПО должны быть в сравнительно небольшом диапазоне в окрестности 0. Возможно, это является причиной фазового сброса (рисета), экспериментально наблюдаемого в мозге при некоторых экспериментальных условиях.

3. Высокая крутизна функции взаимодействия  $f$ , что обеспечивает эффективное влияние ПО на ЦО.

4. Достаточно медленная адаптация амплитуд ПО по сравнению со скоростью фазовой синхронизации между ПО и ЦО.

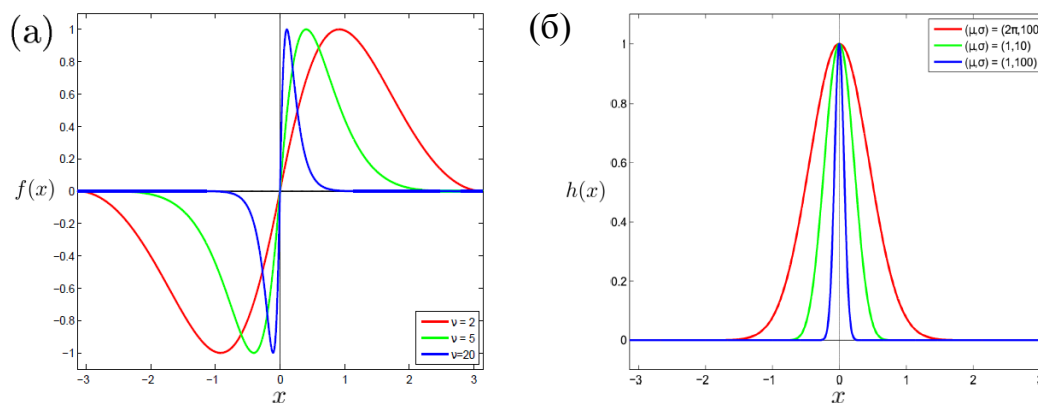


Рис. 4. Вид функций (11): (а) Функция взаимодействия  $f(x)$  при различных значениях  $\nu$ ; (б) функция управления резонансом  $h(x)$  для различных значений  $\mu, \sigma$ .

Следует также отметить, что фазово-осцилляторная модель создает основу для разработки устройств (например, лазерных или на джозефсоновских контактах), реализующих процедуру ППВ в аналоговой форме.

### Стационары в сети фазовых осцилляторов Хонг-Строгаца

В работах [Hong and Strogatz // Phys. Rev. E, 2011; Hong and Strogatz // Phys. Rev. Lett., 2011] была предложена модель, которую можно рассматривать как обобщение модели с ЦО при синхронизирующем воздействии от ПО к ЦО и десинхронизирующем воздействии от ЦО к ПО. Эта модель может представлять интерес для нейрофизиологической интерпретации в том случае, когда центральный управляющий элемент представляет собой не один осциллятор, а ансамбль взаимодействующих осцилляторов.

Модель Хонг-Строгаца состоит из двух популяций фазовых осцилляторов. Все осцилляторы делятся на две группы. Осцилляторы первой группы, называемые конформистами, получают только синхронизирующие воздействия, а осцилляторы второй группы, нонконформисты, получают только десинхронизирующие воздействия. В результате такой организации взаимодействия между осцилляторами конформисты стремятся синхронизоваться с осцилляторами, которые оказывают на них воздействие, а нонконформисты "предпочитают" работать с противофазе с воздействующими на них осцилляторами.

Модель Хонг-Строгаца описывается следующим образом. Пусть  $J$  – набор осцилляторов,  $|J| = n$ ,  $J_1, J_2$  – группы, соответственно, конформистов и нонконформистов,  $|J_1| = n_1$ ,  $|J_2| = n_2$ ,  $n_1 + n_2 = n$ . Динамика модели задается уравнениями

$$\frac{d\theta_i}{dt} = \omega + \frac{k_1}{n} \sum_{j=1}^n \sin(\theta_j - \theta_i), \quad i \in J_1, \quad (12)$$

$$\frac{d\theta_i}{dt} = \omega + \frac{k_2}{n} \sum_{j=1}^n \sin(\theta_j - \theta_i), \quad i \in J_2, \quad (13)$$

где  $\theta_i$  – фазовые переменные,  $\omega$  – собственная частота (одинаковая для всех осцилляторов),  $k_1, k_2$  – параметры связи,  $k_1 > 0$ ,  $k_2 < 0$ .

Обозначим  $k = k_2 / k_1$  и  $\phi_i = \theta_1 - \theta_{i+1}$ ,  $i = 1, \dots, n-1$ .

Упорядочим осцилляторы так, чтобы конформисты имели номера  $1, \dots, n_1$ , а нонконформисты имели номера  $n_1 + 1, \dots, n$ . Перейдя к переменным  $\phi_i$  и перемасштабировав время, уравнения (12)-(13) можно переписать в виде

$$\frac{d\phi_i}{dt} = -\sin\phi_i - \sum_{j=1}^{n-1} \sin\phi_j - \sum_{j=1}^{n-1} \sin(\phi_i - \phi_j), \quad i = 1, \dots, n_1 - 1, \quad (14)$$

$$\frac{d\phi_i}{dt} = -k\sin\phi_i - \sum_{j=1}^{n-1} \sin\phi_j - k \sum_{j=1}^{n-1} \sin(\phi_i - \phi_j), \quad i = n_1, \dots, n - 1, \quad (15)$$

Стационарные точки уравнений (14)-(15) в координатах  $\theta_i$  удовлетворяют уравнениям

$$g_1(\theta_1, \dots, \theta_n) - g_i(\theta_1, \dots, \theta_n) = 0, \quad i = 2, \dots, n_1 - 1, \quad (16)$$

$$g_1(\theta_1, \dots, \theta_n) - kg_i(\theta_1, \dots, \theta_n) = 0, \quad i = n_1 + 1, \dots, n, \quad (17)$$

где

$$g_i(\theta_1, \dots, \theta_n) = \sum_{j=1}^n \sin(\theta_i - \theta_j).$$

**Теорема 1.** Для  $k \neq 0, k \neq -n_2 / n_1$  точка  $(\theta_1, \dots, \theta_n)$  удовлетворяет уравнениям (16)-(17), если и только если выполняется одно из следующих условий

A.  $\sum_{j=1}^n e^{i\theta_j} = 0$ ,

B.  $\theta_i - \theta_j \in \{0, \pi\}$ ,  $i, j = 1, \dots, n$ .

Доказательство теоремы приведено в диссертации.

Теорема 1 является исходным пунктом бифуркационного анализа модели Хонг-Строгаца. Этот анализ был проведен в работе [Burylko et al. // Phys. Rev. E, 2014] и позволил дать описание условий устойчивости стационарных точек.

### Модель селективного внимания

В данном разделе рассматривается модель селективного внимания в виде сети фазовых осцилляторов с центральным элементом и постоянными

синхронизирующими силами связи между осцилляторами. Внешние стимулы представляются либо отдельными ПО, либо наборами (ансамблями) ПО с близкими собственными частотами осцилляторов внутри каждого ансамбля. В биологических терминах ПО представляет колонку неокортекса, кодирующую определенный признак стимула, а ЦО представляет центральный управляющий элемент системы внимания.

Пусть есть два стимула, один из которых захватывается в фокус внимания путем частичной синхронизации соответствующего ему ансамбля ПО с ЦО. Второй стимул является дистрактором (отвлекающим стимулом). Соответствующие ему ПО десинхронизованы по отношению к ЦО. Нас будут интересовать динамические режимы, которые могут возникнуть в такой модели и которые можно интерпретировать как различные режимы формирования фокуса внимания.

Модель внимания описывается уравнениями

$$\frac{d\theta_0}{dt} = \omega_0 + \frac{\alpha}{n} \sum_{i=1}^n \sin(\theta_i - \theta_0) + \frac{\beta}{n} \sum_{j=1}^n \sin(\theta_j - \theta_0), \quad (18)$$

$$\frac{d\theta_i}{dt} = \omega_i + \alpha \sin(\theta_0 - \theta_i), \quad i = 1, \dots, n, \quad (19)$$

$$\frac{d\theta_j}{dt} = \omega_j + \beta \sin(\theta_0 - \theta_j), \quad j = 1, \dots, n, \quad (20)$$

где  $\theta_i$  – фазы осцилляторов,  $\omega_i$  – собственные частоты,  $\alpha, \beta$  – параметры взаимодействия ( $\alpha, \beta \geq 0$ ). Обозначим  $A$  и  $B$  наборы ПО из уравнений (19) и (20), соответственно. Будем предполагать, что собственные частоты ПО в этих группах образуют два непересекающихся интервала длины  $l$ .

Опишем различные динамические режимы, возникающие в модели (18)-(20) в зависимости от значений параметров. Будем рассматривать три следующих типа динамики:

- **GS** (global synchronization – полная синхронизация) – все осцилляторы работают с одинаковой частотой (фокус внимания распространяется на оба стимула);
- **PS<sub>A</sub>, PS<sub>B</sub>** (partial synchronization – частичная синхронизация) – модель работает в режиме частичной синхронизации с одним из ансамблей  $A$  или  $B$  (в фокус внимания включен один из стимулов, кодирующихся, соответственно, активностью ансамблей  $A$  или  $B$ );
- **NS** (no synchronization – отсутствие синхронизации) – текущие частоты хотя бы некоторых ПО в ансамблях  $A$  и  $B$  значительно отличаются от текущей частоты ЦО (фокус внимания не сформировался).

На рис. 5 показаны диаграммы различных режимов синхронизации при  $n \rightarrow \infty$ . Непрерывные линии границ между областями получены аналитически, а кружками и звездочками показаны результаты компьютерных вычислений. Видно, что аналитически вычисленные границы между областями частичной синхронизации и областью отсутствия синхронизации хорошо согласуются с результатами, полученными с помощью компьютерных вычислений.

Интересный феномен заключается в том, что уменьшение силы взаимодействия между ЦО и ПО, кодирующими стимул, находящийся в фокусе внимания, может привести не к фокусировке внимания на другом стимуле, а к разрушению фокуса внимания. В этом случае граница между **GS** и **NS** имеет такую форму, что переход между этими областями может быть осуществлен за счет изменения лишь одного из параметров  $a$  или  $\beta$ .

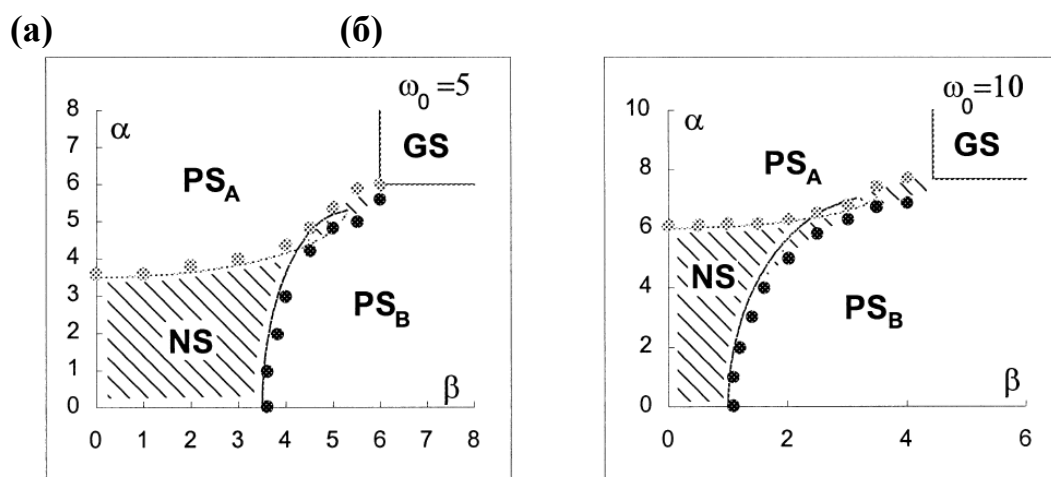


Рис. 5. Области параметров взаимодействия, соответствующие различным режимам синхронизации. Аналитические результаты аппроксимации границ областей показаны сплошными линиями. Кружками и звездочками показаны границы, полученные путем компьютерного моделирования динамики сети. Параметры:  $n = 50$ ,  $\omega_A = 0$ ,  $\omega_B = 10$ ,  $l = 1$ .

### Модель детекции новизны в гиппокампе

Детекцию новизны можно рассматривать как дифференциальный ответ различных структур мозга на поступающий стимул в зависимости от наличия или отсутствия памяти о нем, сохраненной в предшествующем опыте. В дальнейшем мы ограничимся специальным случаем детекции новизны, связанным с ориентационным рефлексом, при котором имеет место изменение активности в гиппокампе [Sokolov // *Neuronal mechanisms of the orienting reflex*, 1975; Vinogradova // *Progr. Neurobiol.*, 1995]: длительная (тоническая) активность на частоте тета-ритма, появляющаяся в гиппокампе при восприятии нового или важного стимула, сменяется кратковременной (фазической) реакцией при повторных предъявлениях того же стимула.

Наша модель следует идее о том, что запоминание информации в гиппокампе управляется фазовыми соотношениями между входными сигналами, поступающими на гиппокамп от энторинальной коры и септума. Рассматривается многомерный входной сигнал (число входов больше двух) со случайными фазовыми сдвигами между отдельными входами. Это приводит к разреженному распределению областей высокой активности в сети при каждой стимуляции и, как следствие, к разреженному кодированию стимулов в памяти сети.

Память в модели возникает как результат преобразования частотно-временного кода на входе сети в пространственно-временной код в памяти сети.

Дополнительно к сказанному, модель имеет следующие новые особенности:

- память формируется не путем адаптации связей между осцилляторами, а путем подходящей модификации собственных частот осцилляторов;
- вспоминание запомненной информации реализуется в виде резонанса осцилляторной активности сети в ответ на новый стимул.

Осцилляторными элементами в нашей модели являются не отдельные нейроны, а ансамбли взаимодействующих возбуждающих и тормозных популяций. Эти элементы представлены упрощенно в виде фазовых осцилляторов. Амплитуда осциллятора представляет среднюю активность ансамбля возбуждающих нейронов. Взаимодействие между осцилляторами описывается в терминах фазовой синхронизации. Осциллятор описывается тремя переменными: текущей фазой, собственной частотой и амплитудой.

Модель детекции новизны представляет собой сеть обобщенных фазовых осцилляторов с тормозной петлей обратной связи (рис. 6). Осцилляторы собраны в группы  $G_j (j = 1, \dots, m)$ , в каждой группе имеется  $q$  осцилляторов. Осцилляторы внутри группы связаны глобальными связями.

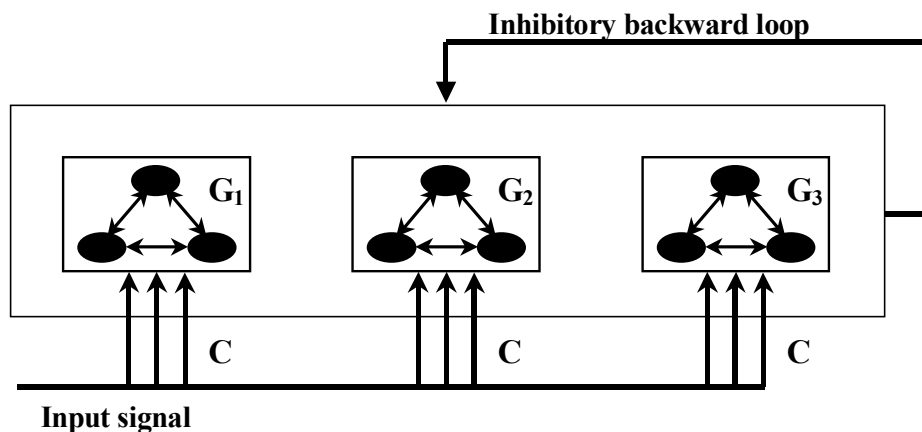


Рис. 6. Архитектура модели детекции. Осцилляторы собраны в группы со связями все-на-всех внутри каждой группы. Связи между группами отсутствуют. Каждый осциллятор получает многомерный входной сигнал  $C$ . Петля тормозной обратной связи предназначена для подавления активности в сети после того, как число осцилляторов в состоянии резонанса превысит пороговое значение.

Входной  $n$ -мерный сигнал  $C = (C_1, \dots, C_n)$  с компонентами  $C_i = \sin(2\pi\omega_0 t + \psi_{ij})$  ( $i = 1, \dots, n, j = 1, \dots, m$ ) кодирует информацию о предъявленном стимуле и представляет собой периодические колебания с частотой  $\omega_0 \in (\omega_{min}, \omega_{max})$  и разными фазовыми сдвигами  $\psi^{ij} \in (-\tau, \tau)$ , подающиеся одновременно на все осцилляторы сети. Фазовые сдвиги имитируют различные временные задержки при передаче сигнала на различные группы осцилляторов. В начальный момент времени (до предъявления каких-

либо стимулов) каждая группа  $G_j$  содержит осцилляторы, собственные частоты которых равномерно распределены в диапазоне  $(\omega_{min}, \omega_{max})$ . По мере предъявления и запоминания стимулов собственные частоты осцилляторов адаптируются в зависимости от параметров  $\omega_0, \psi^{ij}$  входных сигналов.

Основное предположение в модели – каждый осциллятор достигает высокого уровня активности и сохраняет его в течение этапа запоминания стимула, если сигналы, поступающие на осциллятор, приходят синфазно с его текущей фазой (в частности, величины  $\psi^{ij}$  приблизительно одинаковы для заданной группы  $j$ ). При запоминании стимула резонансные осцилляторы постепенно адаптируют свои собственные частоты, формируя кластер с одинаковыми собственными частотами. Если стимул был запомнен, то резонансная активность в сети возникает существенно быстрее, чем когда этот стимул является новым. Благодаря случайному выбору значений  $\psi^{ij}$  для каждого стимула, предъявление стимула приводит к резонансной активности только в небольшом числе случайно выбранных групп осцилляторов, а именно там, где имеет место подходящее совпадение фаз входных сигналов. Такой вид активности в модели воспроизводится как при запоминании, так и при вспоминании (принятии решения о новизне стимула).

Тормозная петля обратной связи предназначена для того, чтобы останавливать активность в сети и адаптацию параметров, когда память на данный стимул уже сформировалась. В сессиях повторных стимуляций одним и тем же стимулом популяция резонансных осцилляторов постепенно увеличивается. Мы предполагаем, что формирование памяти на заданный стимул заканчивается тогда, когда размер популяции осцилляторов в резонансном состоянии превосходит заданный порог  $H$ . Как только это происходит, вся активность в сети подавляется с помощью петли тормозной обратной связи на весь промежуток времени вплоть до начала новой стимуляции.

Модель обучается запоминанию путем предъявления последовательности стимулов. Предъявление каждого стимула ведет к появлению осцилляторной активности в сети. Собственные частоты осцилляторов, установившиеся на предыдущем этапе запоминания, используются как начальные значения собственных частот при новой стимуляции (временного затухания памяти в модели нет). Модифицированные собственные частоты – единственная информация о стимулах, запоминаемая моделью.

На рис. 7 показано поведение числа резонансных осцилляторов в модели для 4 стимулов. Пунктирные вертикальные линии показывают критические моменты, когда происходит дискриминация между новыми и известными стимулами. Из рисунка видно, что число резонансных осцилляторов возрастает от одной стимуляции к другой и впервые преодолевает порог  $Threshold$  во время второй или третьей стимуляции одним и тем же стимулом. Как только время для достижения порога оказывается меньше критического, стимул распознается как известный. В приведенных примерах число повторений

одного и того же стимула, прежде чем он распознается как известный, равно, соответственно, 5, 4, 3 и 4.

В диссертации приведены данные относительно надежности детекции новизны при заполнении памяти модели, показывающие ее работоспособность в практических условиях запоминания в рабочей памяти гиппокампа.

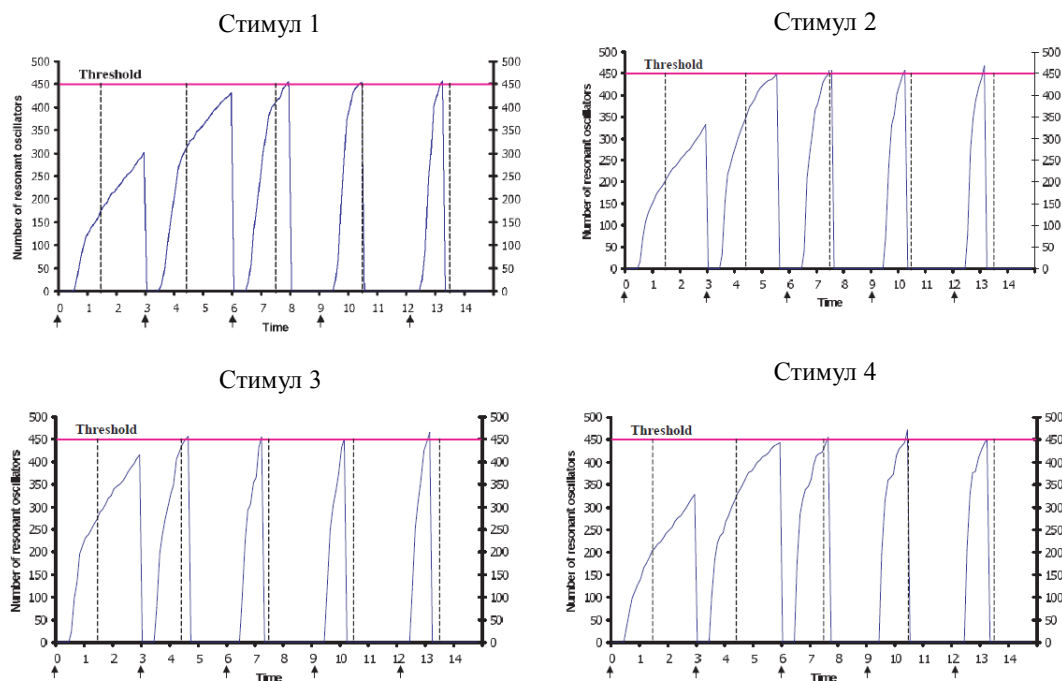


Рис. 7. Графики числа резонансных осцилляторов в сети при стимуляции 4 стимулами. Каждый стимул повторяется 5 раз. Моменты начала стимуляции показаны стрелками. Вертикальные пунктирные линии показывают критические моменты времени, установленные на 1,5 условных единиц времени после начала стимуляции. Горизонтальная линия (Threshold) на уровне 450 показывает порог  $H$ . Стимул считается известным, если число резонансных осцилляторов превышает порог до наступления критических моментов времени. Time – время, Number of resonant oscillators – число резонансных осцилляторов, Threshold – порог.

Отметим, что предложенная модель детекции новизны существенно отличается от известных моделей ассоциативной памяти как своей конструкцией, так и принципами функционирования. Предложенный нами подход основан на использовании ОНС с адаптирующимися значениями собственных частот осцилляторов в результате фазовой синхронизации осцилляторов с многомерным осцилляторным входным сигналом. Резонанс осцилляторов с входными сигналами играет важную роль в нашей модели детекции новизны. Быстрое и одновременное возрастание активности критического числа осцилляторов приводит в действие тормозную обратную связь. Таким образом, сеть отвечает длительной (тонической) реакцией на новые стимулы и короткой (фазической) реакцией на известные стимулы.

## **Модель последовательного выбора объектов на изображении**

Традиционный подход к распознаванию зрительных образов основан на предположении, что зрительные объекты поступают на распознающую систему по одному в определенной последовательности, однако реальные биологические зрительные системы вынуждены работать в условиях, когда зрительные сцены содержат одновременно несколько объектов. Многочисленные экспериментальные исследования показывают, что в последовательном выборе в сложной зрительной сцене задействовано сразу несколько когнитивных функций. Во-первых, интеграция признаков (байндинг) и внимание необходимы, чтобы правильно скомбинировать признаки объектов и отделить признаки одного объекта от признаков других объектов, присутствующих в зрительной сцене. Во-вторых, необходимы детекция новизны и память, чтобы не терять бесполезно время на обработку информации об известных объектах.

В данном разделе будет описана полномасштабная модель для последовательного выбора зрительных объектов, которая совмещает такие когнитивные функции, как интеграция признаков, объектно-ориентированное внимание и детекция новизны. Поскольку мы полагаем, что мозг "не изобретает" специальные механизмы обработки информации для реализации каждой когнитивной функции, но адаптирует сходные механизмы к различным типам обработки, основной задачей при разработке модели было использование небольшого числа общих принципов.

Модель представляет собой иерархию интерактивных модулей, которые ассоциируются с различными этапами обработки зрительной информации. Каждый модуль представляет собой систему интерактивных обобщенных фазовых осцилляторов с синхронизирующими и десинхронизирующими взаимодействиями.

Модель работает с изображениями, представленными в виде нескольких непересекающихся полутоновых объектов. Обработка изображения включает в себя следующие операции.

1. Сегрегация информации относительно различных объектов в соответствии с их пространственной связностью.
2. Последовательный выбор объектов зрительного поля в фокус внимания.
3. Определение признаков объектов, инвариантных по отношению к положению и шкалированию объектов.
4. Представление объектов в рабочей памяти.
5. Детекция новизны объектов (различный ответ выходного слоя на известные и новые объекты).

Функционирование модели основано на принципах синхронизации (фазовой автоподстройки) и резонансного возрастания активности определенных нейронных популяций (осцилляторов).

Внимание реализуется в модели в результате синхронизации ЦО ансамблем осцилляторов, представляющих объект на изображении. Предполагается, что осцилляторы, работающие синхронно с ЦО, формируют фокус внимания. Благодаря резонансу с ЦО, амплитуды осцилляторов в фокусе



внимания возрастают, в то время как амплитуды осцилляторов за пределами фокуса внимания падают до низкого уровня.

Тот же принцип синхронизации и резонанса нейронной активности используется при реализации детекции новизны. Детекция новизны основана на разреженном кодировании объектов группами резонансных осцилляторов в модуле рабочей памяти подобно тому, как это описывалось в предыдущем разделе, с тоническим (длительным) ответом на новые объекты и коротким (фазическим) ответом на известные объекты.

Модель состоит из четырех слоев (модулей) осцилляторов со связями снизу вверх между слоями. Слои сети носят названия *слой выбора объектов* (СВО), *слой локальных признаков* (СЛП), *слой инвариантных признаков* (СИП) и *слой детекции новизны* (СДН). Имеется также дополнительный центральный осциллятор с глобальными прямыми и обратными связями на СВО.

СВО ответствен за группировку информации, поступающей от внешнего входа, в отдельные кластеры в соответствии со связностью объектов. Этот слой также используется для последовательного выбора различных объектов в фокус внимания, используя принцип ППВ. СВО имеет такую же двухмерную решетчатую структуру, как и зрительное поле, так что имеет место взаимнооднозначное соответствие между пикселями сетчатки зрительного поля и осцилляторами СВО. Группировка пикселей в представления объектов реализуется с помощью синхронизирующих локальных связей между осцилляторами в СВО.

СЛП отвечает за преобразование информации об объекте из пиксельного представления в представление в терминах локальных признаков. Эти признаки дают сжатое и специфическое для каждого объекта описание визуальной информации, которая зависит от локальных геометрических характеристик объектов в зрительном поле. Поскольку в СЛП сигналы за пределами фокуса внимания игнорируются, это предотвращает ложные сочетания признаков различных объектов в процессе запоминания.

СИП работает как счетчик признаков, благодаря чему достигается инвариантное кодирование объектов по отношению к положению и масштабу.

СДН предназначен для запоминания объектов в рабочей памяти и принятия решений относительно новизны объектов. Поскольку одни и те же признаки могут появиться в представлении различных объектов, для усиления различий в представлении объектов требуется еще одно преобразование информации. Это осуществляется в СДН, где каждый объект представляется разреженным ансамблем осцилляторов, который специфическим образом связан с объектом в фокусе внимания. СДН с небольшими модификациями повторяет нейронную сеть для детекции новизны, описанную в предыдущем разделе.

На рис. 8 проиллюстрированы принципы работы модели на простом примере с двумя входными изображениями, предъявляемыми одно за другим. Оба изображения содержат по 5 объектов, представляющих собой черно-белые изображения букв двух слов, соответственно, "HELLO" и "WORLD". Каждое изображение обрабатывалось последовательно объект за объектом с

запоминанием и детекцией новизны всех объектов. Порядок, в котором выбирались объекты на изображениях, был случайным.

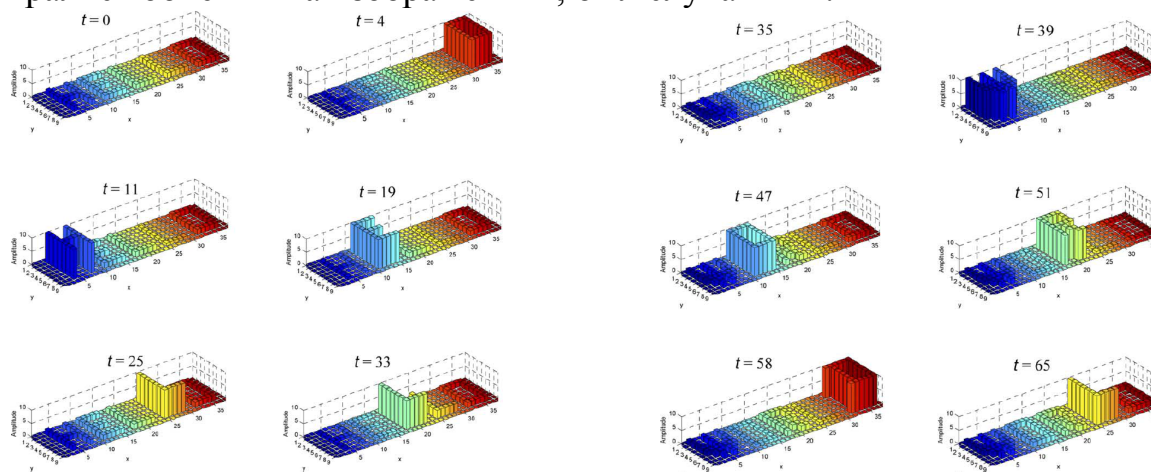


Рис. 8. Динамика амплитуд в СВО: Начальные моменты предъявления изображений слов HELLO и WORLD равны, соответственно, 0 и 35.

Поскольку буквы L и O встречаются на изображениях несколько раз (соответственно, 3 и 2 раза), они воспринимались как известные при втором и третьем попадании в фокус внимания. Другие буквы встречаются на изображениях лишь однажды, поэтому они идентифицируются как новые объекты.

Для новых объектов время реакции варьировало в пределах от 4,1 до 6,0 условных единиц, в то время как для известных объектов это время сокращалось до 1,4-1,8 условных единиц.

### Модель сегментации цветных изображений

Проблема выделения объектов на изображении тесно связана с задачей сегментации изображений: выделенный объект должен быть обособлен от других объектов на изображении и от фона. Сегментация изображения не представляет особого труда, если речь идет об изолированных объектах, расположенных на однородном фоне, существенно отличающемся по своим оптическим характеристикам от выделяемых объектов. На реальных изображениях в типичных случаях объекты пересекаются, а фон обладает значительной изменчивостью.

Нашей целью было разработать биологически правдоподобную модель обработки зрительной информации с дальнейшим ее использованием для управления движениями рук робота. Предлагаемая модель способна выполнять следующие операции: выбор объектов из зрительной сцены, выделение контуров и сегментацию объектов с подавлением шумовых и ложных компонентов изображения. Модель состоит из трех модулей. Модуль селективного внимания сканирует зрительную сцену и выбирает область потенциального интереса, которая может включать один или несколько объектов. Два других модуля используются для детального анализа этой области и сегментации объектов. Область потенциального интереса является входной информацией для модуля определения контуров. Модуль сегментации

объектов в качестве входной информации получает как характеристики пикселей в области потенциального интереса, так и информацию о контурах в этой области. Результатом его работы является последовательная фокусировка внимания на различных объектах с фильтрацией шумов и ложных вкраплений.

Основной интерес в модели сегментации представляет работа модуля сегментации объектов. Этот модуль состоит из двух слоев ПО  $L_1$ ,  $L_2$ , активность которых управляется с помощью ЦО. Первый (нижний) слой осцилляторов реализует синхронизацию ансамбля ПО, представляющего некоторый объект, с ЦО. При этом контур объекта, полученный на предыдущем этапе с помощью модуля выделения контуров, используется для ограничения распространения синхронизации за пределами ансамбля ПО, представляющего данный объект. Второй (верхний) слой ПО преобразует изображение, элиминируя шумы и погрешности в задании контура.



Функционирование модели продемонстрируем на примере изображения, полученного с камеры робота, тренируемого на захват определенных предметов с поверхности. На изображении имеются 4 шара разного цвета, которые являются объектами манипуляций робота. Робот должен обладать способностью взять своей рукой шар заданного цвета. Для этого зрительная система робота должна давать информацию о расположении и границе шара на систему, управляющую движением руки. Хотя представленный пример относится к прикладной робототехнике, используемые методы универсальны и пригодны для любых цветных изображений, в которых целевой объект достаточно значительно отличается по цвету от своего окружения.

Для иллюстрации функционирования модуля селективного внимания на рис. 9 показан фрагмент изображения, представляющий собой область потенциального интереса, в которой находятся потенциальные целевые объекты. Этот фрагмент формируется модулем селективного внимания и служит входной информацией для модуля сегментации объектов.

Контурные на цветном изображении выделяются отдельно по каждому из основных цветов спектра изображения, а затем комбинируются в единые представления контуров. В силу дифференциального характера операторов получения контурных точек, кроме самих контуров на изображении присутствуют шумы и небольшие ложные объекты (рис. 10).

Процесс сегментации зеленого шара показан на рис. 11. Верхний и нижний ряды панелей показывают динамику фаз, соответственно, в слоях  $L_1$  и  $L_2$ . На каждой панели оттенками серого показана разность фаз между ПО и ЦО. Эти величины шкалированы в диапазоне (0, 256), так что чем меньше разность фаз, тем темнее цвет пикселя на рис. 11.

Процесс синхронизации начинается с небольшого квадрата  $S$ , который хорошо виден на рисунке в этот момент времени. Постепенно синхронизация распространяется на весь объект, который должен быть выбран в фокус внимания. Параллельно фазы ПО за пределами зеленого шара стремятся отличаться от фазы ЦО (показано в виде светло-серого фона). Шум, явно присутствующий в слое  $L_1$ , подавляется в слое  $L_2$ .

	
<p>Рис. 9. Фрагмент (область потенциального интереса) исходного изображения.</p>	<p>Рис. 10. Результат выделения контуров на фрагменте изображения.</p>

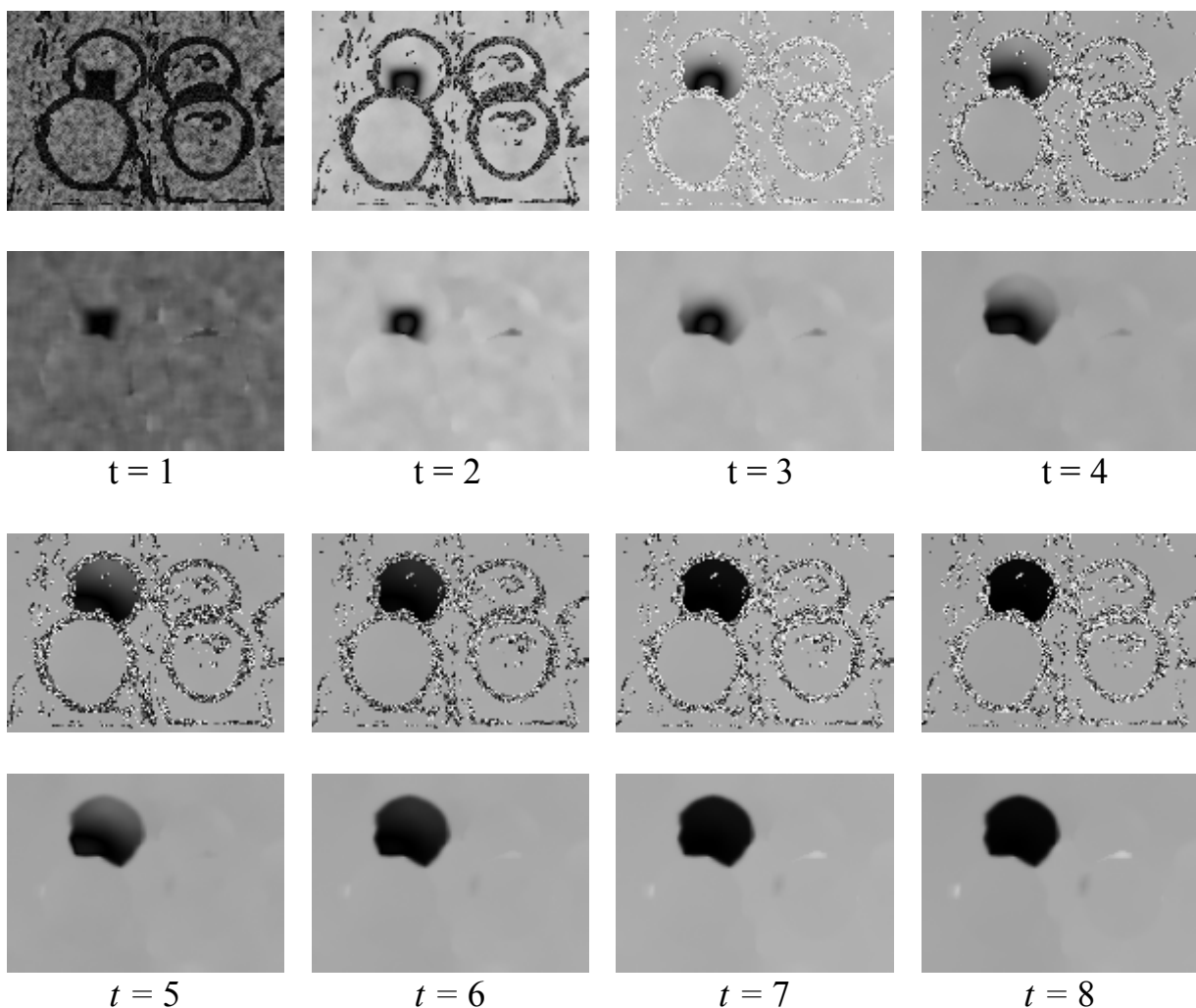


Рис. 11. Сегментация зеленого шара. На каждой панели оттенками серого показана разность фаз между ПО и ЦО. Эти величины шкалированы в диапазоне (0, 256), так что чем меньше разность фаз, тем темнее цвет пикселя.

Видно, что в окончательном изображении границы зеленого шара дефекты исходного контура сглажены. Это достигнуто за счет локального взаимодействия осцилляторов в  $L_2$ . Процесс сегментации других шаров происходит аналогично. Отметим, что процесс сегментации занимает примерно

8-10 условных единиц времени, что соответствует 130-160 мс. Это соответствует экспериментально регистрируемым периодам времени, необходимого для фокусировки внимания.

### Модель зрительного поиска

Зрительный поиск – это определенный тип решения задачи восприятия, которая включает активное зрительное исследование окружающей среды с целью найти определенный объект (целевой объект), находящийся в окружении других объектов (объектов-дистракторов).

Уже в ранних психофизических экспериментах [Treisman and Gelade // Cogn. Psychol., 1980] было показано, что задачи зрительного поиска могут быть разной степени трудности в зависимости от того, насколько целевой объект заметен среди дистракторов. Трудность задачи отражается в промежутке времени, необходимом наблюдателю для решения задачи зрительного поиска и в числе ошибок, которые наблюдатель делает (рис. 12).

*Поиск по признаку* соответствует случаю, когда целевой объект и дистракторы ясно различаются по одному признаку: цвету, форме, ориентации, размеру и т.д. В этом случае время реакции (reaction time - RT) мало (около 300 мс) и не зависит от числа объектов  $n$ , одновременно присутствующих на дисплее (график RT( $n$ ) почти параллелен оси абсцисс).

*Поиск по сочетанию признаков* имеет место, когда целевой объект имеет какой-то признак общий с дистракторами. В этом случае время реакции увеличивается (более 500 мс), и RT( $n$ ) заметно возрастает с ростом  $n$ .

*Поиск по пространственной форме* – это задача еще более высокой сложности. Примером такой задачи является поиск цифры 2 среди цифр 5. В этом случае реакция может занять несколько секунд, а график RT( $n$ ) становится намного более крутым, чем при поиске по сочетанию признаков.

На рис. 13 приведены графики RT( $n$ ) для трех задач, приведенных на рис. 12, в случае, когда изображение содержит целевой объект (сплошные линии) и не содержит целевой объект (пунктирные линии).

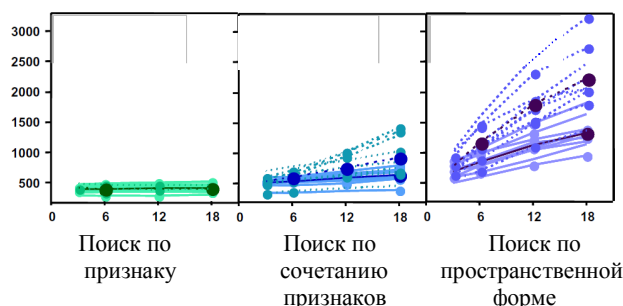
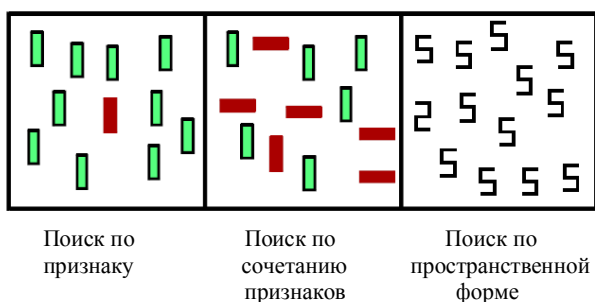


Рис. 12. Примеры задач зрительного поиска разной степени сложности.

Рис. 13. Средние времена реакции для задач зрительного поиска. Сплошные линии – при наличии целевого объекта, пунктирные линии - при отсутствии целевого объекта [Wolfe et al. // Vision Res., 2010].

В качестве модели зрительного поиска мы использовали осцилляторную версию процедуры ППВ. Это сеть с радиальной архитектурой, построенная из обобщенных фазовых осцилляторов. Каждый зрительный объект представляется в модели одним ПО. Заметность объекта описывается величиной связи от ПО к ЦО. Как и раньше, будем эквивалентно считать эту величину амплитудой колебаний осциллятора.

Наша модель объединяет как параллельные, так и последовательные механизмы обработки информации. Первые используются для выбора объекта в фокус внимания, а вторые - для ограниченного перебора объектов.

Различие в асимптотической динамике между ПО можно выявить, проверяя, к какой границе значений амплитуды ПО стремится амплитуда данного ПО, к верхней или к нижней. Прежде чем делать такую проверку, необходимо подождать некоторое время  $T$ , пока идут переходные процессы.

Зададим два порога,  $H_{high}$  и  $H_{low}$ . Если траектория  $a_i(t)$  пересекает порог  $H_{high}$ , то считаем, что  $i$ -й объект выбран в ФВ.

Массивные компьютерные эксперименты показали, что модель может демонстрировать следующие типы динамики (рис.14).

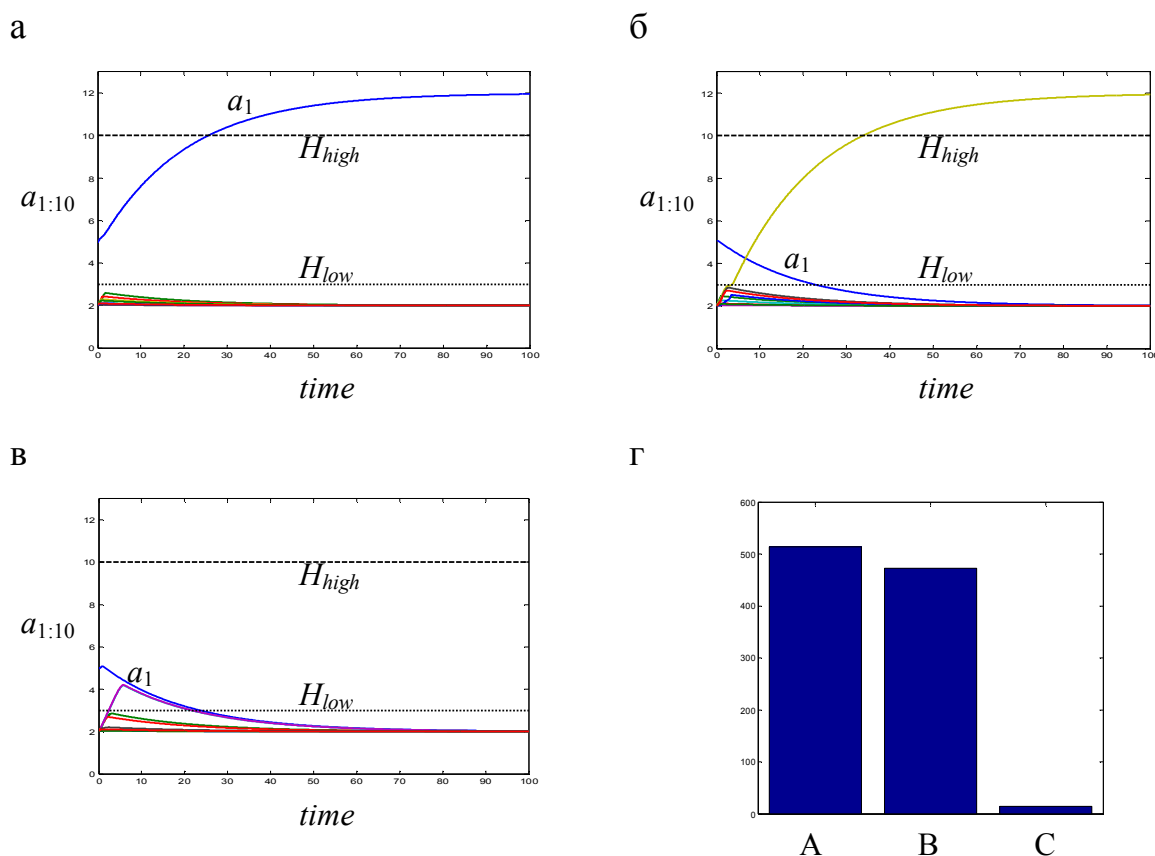


Рис. 14. Типы динамик модели: (а) выбор целевого объекта; (б) выбор объекта-дистрактора; (в) фокус внимания не сформировался; (г) гистограмма частоты возникновения различных динамических режимов. Значения параметров:  $n = 10$ ,  $a_1(0) = 5$ ,  $a_i(0) = 2$  ( $i = 2, \dots, 10$ ); time – время в условных единицах.

А.  $a_1(t) > H_{high}$  для  $t > T$ , и "целевой" ПО является единственным осциллятором, для которого это условие выполняется. Это означает, что в фокус внимания был выбран целевой объект (рис. 14а). В этом случае процесс выбора объектов в фокус внимания заканчивается.

В.  $a_i(t) > H_{high}$  для  $t > T$  и для некоторого значения индекса  $i > 1$ . Это означает, что в фокус внимания был выбран объект-дистрактор (рис. 14б). В этом случае выбора объектов в фокус внимания продолжится в других прогонах модели до тех пор, пока не будет сделан правильный выбор целевого объекта.

С.  $a_i(t) < H_{low}$  для всех  $i$  при  $t > T$  (рис. 14в). Это означает, что за время  $T$  фокус внимания не сформировался. В этом случае процесс выбора объектов продолжается на следующих прогонах модели аналогично случаю В.

На рис. 14г приведен пример гистограммы, показывающий частоту возникновения событий А, В, С при проведении 1000 прогонов модели. Отметим, что частота события С всегда невелика, не более 3-5%.

На рис. 15а показаны вероятности  $r_n$  выбора в фокус внимания целевого объекта за один прогон модели.

Графики показывают, что  $r_n$  убывает при увеличении числа объектов  $n$ . При фиксированной начальной амплитуде осцилляторов-"дистракторов" значения  $r_n$  становятся меньше, по мере того как уменьшается начальное значение амплитуды "целевого" осциллятора. В терминах модели это означает, что вероятность выбора целевого объекта в фокус внимания убывает, если заметность целевого объекта падает и его становится труднее различить среди дистракторов.

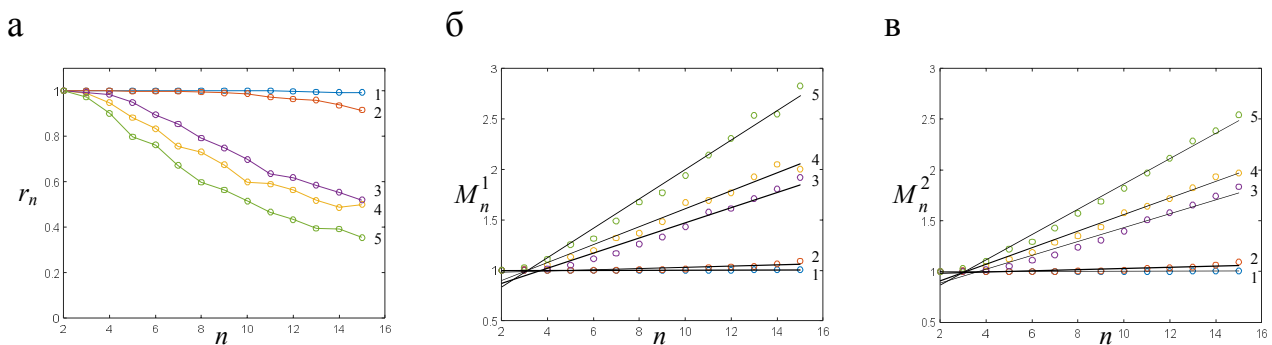


Рис. 15. Результаты численных экспериментов: (а) вероятности  $r_n$  как функции  $n$ ; (б) среднее число прогонов  $M_n^1$ , необходимых для выбора целевого объекта в случае выбора с разрешенным возвратом; (в) среднее число прогонов  $M_n^2$ , необходимое для выбора целевого объекта в случае подавления возврата. Параметры  $a_i(0) = 1$ ,  $i = 2, \dots, n$ . Для всех панелей нумерация линий соответствует следующим значениям  $a_1(0)$ : (1)  $a_1(0) = 12$ , (2)  $a_1(0) = 8$ , (3)  $a_1(0) = 6$ , (4)  $a_1(0) = 4$ , (5)  $a_1(0) = 3$ . Прямые линии на панелях (б), (в) – линии линейной регрессии.

Значение  $r_n$  является скрытым параметром. Его невозможно непосредственно получить в психофизических экспериментах. Однако  $r_n$  можно использовать для вычисления среднего числа попыток  $M$ , необходимых для выбора целевого объекта.

Поскольку среди исследователей нет определенного мнения относительно того, имеет ли место при зрительном поиске подавление возврата, мы рассмотрели обе эти ситуации. Обозначим  $M_n^1$  число попыток поиска, которое необходимо сделать в случае разрешенного возврата (разрешен повторный просмотр объектов), а через  $M_n^2$  - в случае подавления возврата (просмотренные объекты игнорируются).

Результаты вычислений  $M_n^1$  и  $M_n^2$ , приведенные на рис. 15б,в, хорошо соответствуют экспериментальным данным относительно времен реакции в задачах зрительного поиска различной степени трудности (сравни с рис. 12).

Компьютерные вычисления показали, что  $RT(n)$  с ростом  $n$  возрастает линейно, что согласуется с данными психофизических экспериментов [Palmer et al. // J. Exp. Psychol. Hum. Percept. Perform., 2011]. Это не было запланировано заранее и является подтверждением того, что модель верно отражает нейронные механизмы, лежащие в основе работы селективного внимания при зрительном поиске.

### **Модель слежения за несколькими объектами**

Одной из основных экспериментальных парадигм при исследовании объектно-ориентированного внимания является слежение за несколькими объектами (СНО). В стандартно организованных экспериментах по СНО [Pylyshyn and Storm // Spat. Vis., 1988; Pylyshyn // Cognition, 2001; Scholl // Cognition, 2001] наблюдатель (испытуемый) смотрит на дисплей, содержащий  $m$  простых идентичных объектов (до 10-12 объектов, таких как точки, знаки +, круги и т.д.). Определенное число этих объектов коротко подсвечивается с тем, чтобы испытуемый в дальнейшем рассматривал эти объекты как целевые. Остальные объекты являются дистракторами. После этого все объекты начинают двигаться по экрану дисплея независимо друг от друга в случайных направлениях. Движения ограничены двумя условиями: объекты не должны подходить слишком близко друг к другу и не должны выходить за пределы экрана. Цель испытуемого – следить за целевыми объектами без саккад (зрачки глаз фиксируются в центре экрана). В какой-то момент один из объектов подсвечивается вспышкой. Испытуемый должен нажать одну из двух клавиш клавиатуры в зависимости от того, был ли подсвеченный объект целевым или дистрактором.

Было показано, что тренированные испытуемые могут весьма эффективно выполнять задания по СНО. Хотя число ошибок возрастает с увеличением числа целевых объектов, даже в случае 5 целей число правильных ответов может быть на уровне 85%.

Архитектура модели СНО показана на рис. 16. Эта архитектура соответствует случаю слежения за тремя объектами (в общем случае число



слоев в сети совпадает с числом целевых объектов). Каждый из слоев имеет ту же конструкцию, что и модель выбора объектов на изображении с той особенностью, что выбор ограничивается одним объектом. Слои рассматриваются как модули системы внимания. Каждый модуль предназначен для слежения за своим целевым объектом, а взаимодействие между модулями организовано так, чтобы разные модули не следили за одним и тем же объектом.

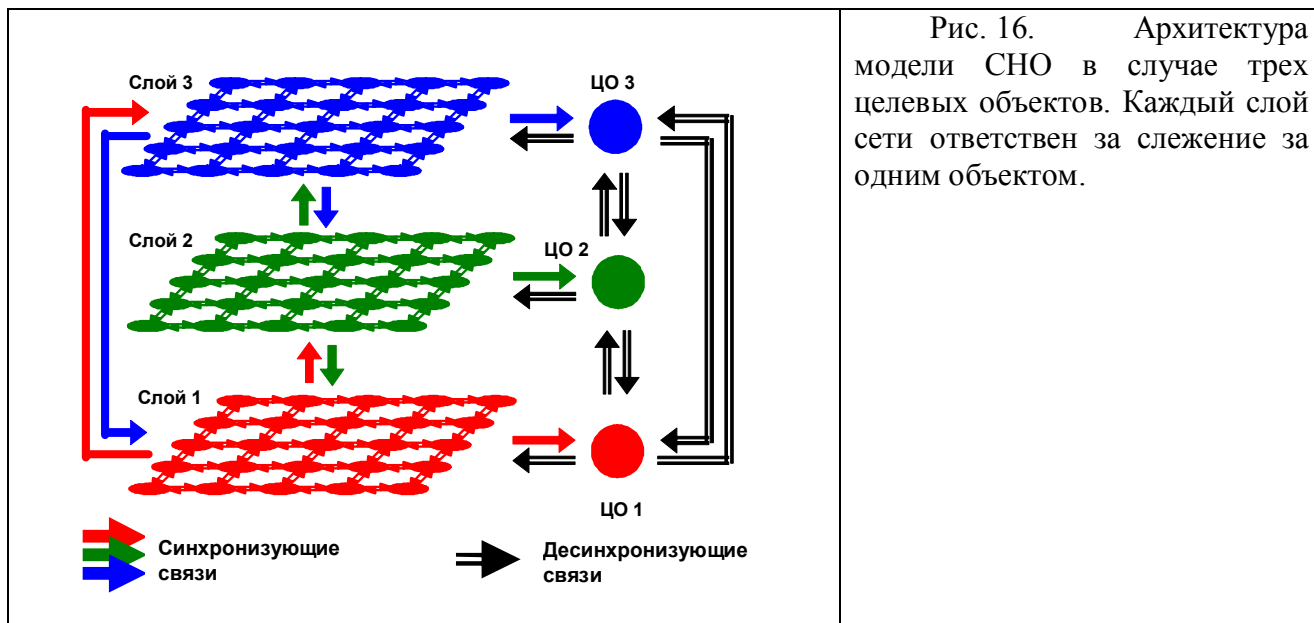


Рис. 16. Архитектура модели СНО в случае трех целевых объектов. Каждый слой сети ответствен за слежение за одним объектом.

Если объекты движутся по зрительному полю достаточно медленно и не пересекаются в процессе перемещений, фокус внимания (после того, как он сформирован) остается достаточно устойчивым в силу резонанса ПО, включенных в фокус внимания. Резонансные осцилляторы намного сильнее влияют на ЦО в своем слое, чем другие осцилляторы, и это предотвращает скачки внимания с одних объектов на другие. Однако, если скорость перемещения объектов становится сравнительно большой относительно скорости процессов синхронизации и резонанса, внимание может спонтанно переключиться с одного объекта на другой. Это приводит к ошибкам при идентификации целевых объектов и объектов-дистракторов.

Для оценки качества функционирования модели проводилось по 50 прогонов вычислений для каждого из значений числа целевых объектов  $k = 2, 3, 4, 5$ . Результаты вычислений представлены на рис. 17 совместно с экспериментальными данными. Модель дает достаточно хорошую аппроксимацию экспериментальных данных. Существенное отклонение имеет место в случае двух целевых объектов. Это, вероятно, связано с тем, что параметры модели были фиксированными для всех случаев слежения, в то время как в реальности не исключена подстройка параметров, зависящая от числа объектов на изображении.

Отметим, что модель равно успешно может применяться, когда объекты в зрительном поле не идентичны или даже меняют свою форму. Это

существенно, например, если движения совершаются в трехмерном пространстве, когда проекция объектов на сетчатку постоянно меняется.

### Модель восприятия неоднозначных изображений

Феномен бистабильности восприятия хорошо известен [Necker // Lond. Edinburgh Philos. Mag. J. Sci., 1832; Rubin // Visuell wahrgenommene Figuren, 1921]. Экспериментально бистабильность обнаруживается при восприятии некоторых изображений, таких как куб Неккера или "ваза" Рубина, или при подаче различных изображений на левый и правый глаза (бинокулярная конкуренция). При такой внешней стимуляции зрительной системы испытуемый ощущает постоянное переключение между различными представлениями одного и того же изображения. В психологических экспериментах с неоднозначными изображениями [Fox and Herrmann // Percept. Psychophys., 1967; Borsellino et al. // Biol. Cybern., 1972] было показано, что для временных интервалом между переключениями восприятия (так называемые времена переключения восприятия (ВПВ, reversal times) характерно гамма-распределение. Этот факт обычно рассматривается как один из основных в теории неоднозначного восприятия, требующий теоретического и модельного объяснения.

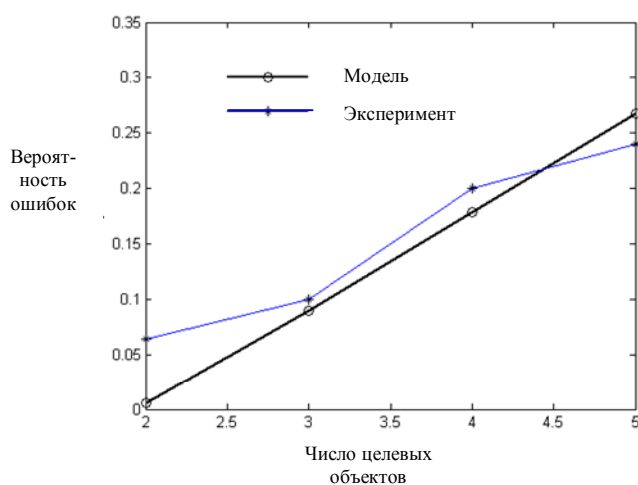


Рис. 17. Вероятность ошибок в модели СНО и в эксперименте. Экспериментальные данные взяты из работы [Oksama and Hyönä // Visual Cogn., 2004].

В качестве модели неоднозначного восприятия мы использовали модель внимания с параметрами, соответствующими режиму NS. При выбранных параметрах происходит спонтанное и хаотическое переключение синхронизации ЦО между двумя ансамблями ПО, что в данном контексте интерпретируется как переключение между двумя восприятиями. Гистограмма времен пребывания системы в разных перцептивных состояниях показана на рис. 18. Непрерывная кривая – аппроксимация гистограммы гамма-распределением. Параметры этого распределения хорошо согласуются с экспериментальными данными (рис. 19).

## Модель запоминания временных последовательностей

Запоминание последовательностей и их вспоминание является непростой задачей, что проявляется в ошибках при вспоминании сложных последовательностей, особенно в случае, когда они коррелированы с другими запомненными последовательностями. Вспоминание может быть улучшено, если последовательность маркирована какой-то меткой, ассоциированной с контекстом, при котором последовательность была запомнена. Мы предполагаем, что такие метки хранятся в верхних областях новой коры, а их взаимодействие с представлениями запомненных объектов осуществляется с помощью синхронизации между фронтальными и ассоциативными зонами, что приводит к модификации связей между этими областями. Мы использовали этот механизм для разработки биологически правдоподобной нейросетевой модели, в которой улучшено вспоминание сложных последовательностей. Эта модель воспроизводит некоторые известные эффекты, связанные с запоминанием и вспоминанием последовательностей.

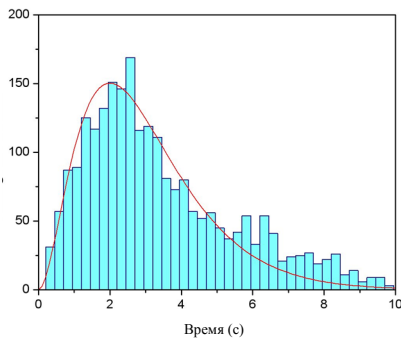


Рис. 18. Гистограмма времен пребывания модели в различных перцептивных состояниях (временные интервалы между переключениями восприятия). Непрерывная кривая — аппроксимация гистограммы гамма-распределением с двумя параметрами.

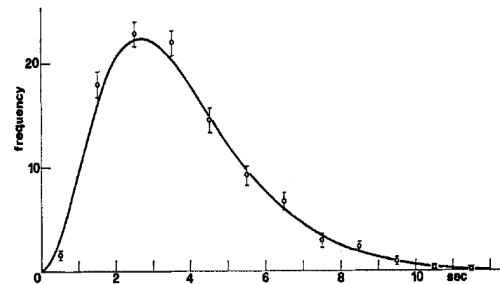


Рис. 19. Экспериментальная гистограмма времен пребывания индивидуума в различных перцептивных состояниях (временные интервалы между переключениями восприятия [Borsellino et al. // Biol. Cybern., 1972]).

Модель построена из импульсных нейронов типа нейронов Ходжкина-Хаксли, организованных в виде двух слоев. Нижний слой кодирует последовательные события в виде активности модулей, которые состоят из возбуждающих и тормозных нейронов со связями все-на-всех. Конвергентные связи идут от нижнего слоя к верхнему. Верхний слой кодирует метки, ассоциированные с различными последовательностями. Обратные связи от верхнего слоя к нижнему реализуют различение последовательностей. Связанная с запоминанием модификация синаптических связей осуществляется по принципу STDP или анти-STDP.

Отметим следующие характеристики нашей модели.

1. Модель способна надежно запоминать и вспоминать несколько последовательностей событий. Для того чтобы начать вспоминание, на вход модели должен быть подан первый (или любой другой) член последовательности, а также метка этой последовательности.

2. Функционирование модели основано на осцилляторной активности в тета-диапазоне. Это позволяет разделить последовательность на отдельные события, представляя каждое событие в своем периоде тета-ритма.
3. Вся выбранная последовательность (или подпоследовательность, начинающаяся с заданного события) будет воспроизведена в том же временном порядке, как события расположены в последовательности. Процесс вспоминания занимает короткое время: вся последовательность проигрывается в течение одного периода тета-ритма.
4. Модель может работать с наборами последовательностей с пересекающимися членами. Сигнал метки от верхнего слоя позволяет модели выбрать правильное продолжение последовательности в случае неоднозначной ситуации.
5. Модель с анти-STDP правилом запоминания позволяет реализовать обратное проигрывание последовательности в соответствии с экспериментальными данными.

## ЗАКЛЮЧЕНИЕ

В данной работе нас интересовала связь когнитивных функций с динамическими аспектами работы мозга. Вопрос этот не праздный, так как многие существующие модели эту связь игнорируют и на функциональном уровне зачастую получают вполне удовлетворительные результаты, иногда даже превосходящие возможности естественного интеллекта, как это уже продемонстрировано в некоторых задачах обучения и распознавания. Тем не менее, многие экспериментальные данные указывают на то, что быстрая смена типов активности и гибкая адаптация мозга являются неотъемлемыми атрибутами его успешного функционирования. Каким образом мозг использует свои динамические режимы? Есть ли общие принципы в архитектурах, которые эти режимы поддерживают, и в механизмах, которые эти режимы реализуют? Какие типы обработки информации могут быть объяснены и воспроизведены на основе колебательной активности и ее синхронизации? Не претендуя на исчерпывающие ответы на эти вопросы, мы полагаем, что разработанные нами модели дают достаточные основания, чтобы утверждать, что синхронизации нейронной активности лежит в основе широкого спектра когнитивных функций. Можно надеяться, что в дальнейшем этот спектр будет расширяться.

Помимо синхронизации, в большинстве наших моделей используются принципы резонанса и адаптации частоты. Эти принципы не следует понимать буквально как какие-то физические процессы, которые могут быть зарегистрированы *in vivo*. Оба этих принципа в терминах моделей, построенных из фазовых осцилляторов или их обобщений, предназначены для того, чтобы отражать синхронизирующее или десинхронизирующее взаимодействие между нейронными популяциями, и не сводятся, например, к изменениям в амплитуде колебаний или к разнице между частотой колебаний в активном и пассивном состоянии. Для того чтобы придать этим характеристикам физический смысл,

требуются детальные модели на уровне нейронов типа Ходжкина-Хаксли. Некоторые шаги в этом направлении мы предприняли, но вопрос о том, насколько фазовые модели могут быть адекватно воспроизведены на детальном уровне, остается пока открытым.

Здесь проходит граница применимости разработанных нами моделей и методов. Они пригодны для решения общих вопросов относительно возможности использования колебательных процессов и синхронизации для реализации различных когнитивных функций, однако в них лишь в самой общей форме использовались как морфологические, так и нейрофизиологические характеристики, связанные с функционированием структур мозга при выполнении когнитивных функций. Редукционизм и предельное упрощение были неизбежны для начальных шагов в понимании механизмов когнитивных функций. Дальнейшее продвижение потребует более детального описания вовлеченных структур и осцилляторных процессов с учетом реальных частотных и фазовых соотношений, а также взаимодействия различных типов колебаний в разных частотных диапазонах. Эта задача требует не только намного больших вычислительных ресурсов, чем те, что были использованы в данной работе, но и больших знаний относительно биологических нейронных сетей, участвующих в различных когнитивных функциях. Наиболее проблемными областями с точки зрения осцилляторного моделирования являются задачи распознавания и ассоциативной памяти. Здесь ОНС пока значительно уступают традиционным коннекционистским подходам, что ограничивает их использование в практически полезных устройствах. Преодоление этого разрыва – один из вызовов для специалистов по ОНС. Нам представляется, что полученные результаты будут служить вектором, задающим направление будущих исследований.

С точки зрения нейронных архитектур в данной работе основной акцент сделан на архитектуры с центральным элементом и их обобщения. По нашему мнению, возможности этого класса архитектур остаются до настоящего времени недооцененными. Наличие хабов в нейронных сетях хорошо известно из исследований анатомических и функциональных связей. Этот факт, однако, редко принимается во внимание в силу традиционных представлений о распределенности представления информации в мозге и желания ограничиться хорошо разработанными методами статистической физики, которые легче всего прилагаются к однородным нейросетевым архитектурам. Между тем, как видно из наших моделей, сети с хабами являются естественным подходом к реализации процедуры ППВ, особенно если эта процедура реализуется на принципах синхронизации.

## **ВЫВОДЫ**

1. Изучение математических и компьютерных моделей осцилляторных нейронных сетей убедительно показывает, что синхронизация нейронной активности является универсальным механизмом, позволяющим мозгу

эффективно обрабатывать информацию при решении разнообразных когнитивных задач.

2. Сети из фазовых осцилляторов и обобщенных фазовых осцилляторов предоставляют эффективный математический аппарат для построения компьютерных моделей когнитивных функций мозга, таких как селективное внимание, детекция новизны, последовательный выбор объектов на изображении, сегментация цветных изображений, зрительный поиск, слежение за несколькими объектами, восприятие неоднозначных изображений. Динамические режимы, возникающие в таких сетях, в частности, режимы синхронизации, естественным образом интерпретируются в терминах решения когнитивных задач.

3. Сети фазовых осцилляторов и обобщенных фазовых осцилляторов с центральным элементом позволяют аналитически исследовать динамические режимы, возникающие в таких сетях, и являются подходящим инструментом для реализации осцилляторной версии процедуры ППВ, встроеной в реализацию многих когнитивных функций.

4. Формирование и переключение фокуса внимания может описываться в терминах синхронизации между информационными и управляющими областями коры.

5. Результаты модельного исследования детекция новизны в гиппокампе позволяют высказать и обосновать гипотезу, что эта когнитивная функция связана с разреженным кодированием и адаптацией собственных частот гиппокампальных осцилляторов, работающих на частоте тета-ритма.

6. Сети фазовых осцилляторов позволяют разработать полномасштабную модель обработки полутонных и цветных изображений, объединяющую различные когнитивные функции в систему, построенную на универсальных принципах.

7. Результаты психофизических экспериментов по зрительному поиску в задачах различной степени трудности могут быть объяснены с помощью модели синхронизации в сети обобщенных фазовых осцилляторов.

8. Ошибки при слежении за движением целевых объектов среди объектов-дистракторов могут быть объяснены как следствие ограниченности фазового пространства, в котором происходит работа системы внимания, если ее описывать в фазово-частотных характеристиках.

9. Гамма-распределение времен восприятия неккерского куба может быть объяснено в терминах частичной синхронизации и ее бистабильности в системе фазовых осцилляторов с центральным элементом.

10. Запоминание и вспоминание последовательности событий может быть реализовано с помощью осцилляторной нейросетевой модели, построенной из импульсных нейронов. Число ошибок при вспоминании запомненной последовательности можно снизить при учете контекста, в котором происходило запоминание.

## Публикации по теме диссертации:

### - статьи в отечественных журналах, входящих в перечень ВАК:

1. Казанович Я.Б. и др. Осцилляторные нейронные сети. Математические результаты и приложения. // Математическое моделирование. 1992. Т. 4, № 1. С. 3-43.
2. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б. Моделирование “предвнимательной” и “внимательной” обработки информации на основе синхронизации нейронной активности. // Радиофизика. 1994. Т. 37, № 8. С. 933-944.
3. Казанович Я.Б., Борисюк Р.М. Синхронизация в нейронной сети фазовых осцилляторов с центральным элементом. // Математическое моделирование. 1994. Т. 6, № 8. С. 45-60.
4. Борисюк Р.М., Борисюк Г.Н., Казанович Я.Б. Моделирование предвнимания и внимания на основе принципа синхронизации. // Биологические мембраны. 1997. Т. 14, № 6. С. 614-620 (перевод на английский: Borisyuk G.N., Borisyuk R.M., and Kazanovich Y.B. Preattention and attention modeling based on synchronization principle. // Membr. Cell Biol. 1998. V. 11(6). P. 753-761).
5. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р. Моделирование динамики нейронной активности и обработка информации в мозге – итоги “десятилетия”. // Успехи физических наук. 2002. Т. 172, № 10. С. 1189-1214 (перевод на английский: Borisyuk G.N., Borisyuk R.M., Kazanovich Y.B., Ivanitskii G.R. Models of neural dynamics in brain information processing – the developments of 'the decade'. // Physics – Uspekhi. 2002. V. 45(10). P. 1073-1095).
6. Чик Д., Борисюк Р.М., Казанович Я.Б. Режимы синхронизация в сети нейронов Ходжкина-Хаксли с центральным элементом. // Математическая биология и биоинформатика. 2008. Т. 3, № 1. С. 16-35.
7. Казанович Я.Б. Теория временной корреляции и модели сегментации зрительной информации в мозге. // Математическая биология и биоинформатика. 2010. Т. 5, № 1. С. 43-97.
8. Казанович Я.Б., Мысин И.Е. Как животные ориентируются в пространстве? Клетки места и клетки решетки. // Математическая биология и биоинформатика. 2015. Т. 10, № 1. С. 88-115.

### - статьи в международных журналах

9. Kazanovich Y.B., Borisyuk R.M. Synchronization in a neural network of phase oscillators with the central element. // Biol. Cybern. 1994. V. 71. P. 177-185.
10. Kazanovich Y.B., Borisyuk R.M. Dynamics of neural networks with a central element. // Neural Networks. 1999. V. 12(3). P. 441-454.
11. Borisyuk R., Borisyuk G., Kazanovich Y. Synchronization of neural activity and information processing. // Behavioral and Brain Science. 1999. V. 21(6). P. 833.
12. Borisyuk R., Denham M., Kazanovich Y., Hoppensteadt F., Vinogradova O. An oscillatory neural network model of sparse distributed memory and novelty detection. // BioSystems. 2000. V. 58. P. 265-272.

13. Borisyuk R., Denham M., Kazanovich Y., Hoppensteadt F., Vinogradova O. Oscillatory model of novelty detection. // *Network: Computation in Neural Systems*. 2001. V. 12. P. 1-20.
14. Kazanovich Y., Borisyuk R. Object selection by an oscillatory neural network. // *BioSystems*. 2002. V. 67(1-3). P. 103-111.
15. Borisyuk R.M., Kazanovich Y.B. Oscillatory neural network model of attention focus formation and control. // *BioSystems*. 2003. V. 71. P. 29-38.
16. Kazanovich Y.B., Borisyuk R.M. Synchronization in oscillator systems with a central element and phase shifts. // *Progr. Theor. Phys.* 2003. V. 110(6). P. 1047-1058.
17. Borisyuk R., Kazanovich Y. Oscillatory model of attention-guided object selection and novelty detection. // *Neural Networks*. 2004. V. 17. P. 899-915.
18. Kazanovich Y.B., Borisyuk R.M. An oscillatory neural model of multiple object tracking. // *Neural Computation*. 2006. V. 18. P. 1413-1440.
19. Borisyuk R.M., Kazanovich Y.B. Oscillations and waves in the models of interactive neural populations. // *BioSystems*. 2006. V. 86 (1-3). P. 53-62.
20. Kazanovich Y.B. Nonlinear dynamics modeling and information processing in the brain. // *Optical Memory & Neural Networks*. 2007. V. 16(3). P. 111-124.
21. Chik D., Borisyuk R., Kazanovich Y. Selective attention model with spiking elements. // *Neural Networks*. 2009. V. 22(7). P. 890-900.
22. Borisyuk R., Kazanovich Y., Chik D., Tikhanoff V., Cangelosi A. A neural model of selective attention and object segmentation in the visual scene: An approach based on partial synchronization and star-like architecture of connections. // *Neural Networks*. 2009. V. 22(5-6). P. 707-719.
23. Borisyuk R., Chik D., Kazanovich Y. Visual perception of ambiguous figures: Synchronization based neural models. // *Biol. Cybern.* 2009. V. 100(6). P. 491-504.
24. Borisyuk R., Chik D., Kazanovich Y. Selective attention model of moving objects. // *Neural Network World*. 2009. V. 19(5). P. 429-445.
25. Burylko O., Kazanovich Y., Borisyuk R. Bifurcations in phase oscillator networks with a central element. // *Physica, D*. 2012. V. 241. P. 1072-1089.
26. Borisyuk R., Chik D., Kazanovich Y., da Silva Gomes J. Spiking neural network model for memorizing sequences with forward and backward recall. // *BioSystems*. 2013. V. 112. P. 214-223.
27. Kazanovich Y., Burylko O., Borisyuk R. Competition for synchronization in a phase oscillator system. // *Physica, D*. 2013. V. 261. P. 114–124.
28. Burylko O., Kazanovich Y., Borisyuk R. Bifurcation study of phase oscillator systems with attractive and repulsive interaction. // *Phys. Rev., E*. 2014. V. 90. P. 022911 (1-18).
29. Kazanovich Y., Borisyuk R. Reaction times in visual search can be explained by a simple model of neural synchronization. // *Neural Networks*. 2017. V. 87. P. 1-7.
30. Burylko O., Kazanovich Y., Borisyuk R. Winner-take-all in a phase oscillator system with adaptation. // *Scientific Reports*. 2018. V. 8. P. 416.