

Нейросетевая модель слежения за несколькими объектами*

Казанович Я.Б.¹, Борисюк Р.М.^{1,2}

¹Институт математических проблем биологии РАН, Пущино, Московская область, yakov_k@impb.psn.ru

²Центр теоретической и вычислительной нейробиологии, Университет г. Плимут, Плимут, Великобритания, rborisyuk@plymouth.ac.uk

Аннотация. Психологические эксперименты показывают, что человек способен одновременно следить за пятью независимо перемещающимися зрительными объектами. В данной работе предлагается нейросетевая модель этого феномена. Модель работает с набором идентичных зрительных объектов, движущихся в заданном поле. На начальном этапе функционирования при неподвижных объектах модель выбирает в фокус внимания подмножество объектов, которые исходно заданы как целевые. Остальные объекты считаются дистракторами, отвлекающими внимание от слежения за целевыми объектами. В дальнейшем при движении объектов модель должна по возможности сохранять начальное разделение объектов на целевые и дистракторы. Идея реализации модели состоит в использовании многослойной осцилляторной сети, в которой каждый слой предназначен для слежения за одним целевым объектом. Будет показано, что одновременное слежение за несколькими объектами может быть реализовано при подходящем выборе синхронизирующих/десинхронизирующих взаимодействий между слоями. Результаты имитационного моделирования показывают, что в случае пересекающихся объектов вероятность ошибки слежения (принятие дистрактора за целевой объект или целевого объекта за дистрактор) возрастает при увеличении числа дистракторов. Такая зависимость имеет место и в психологических экспериментах. Кроме того будет описано поведение модели в случае работы с пересекающимися объектами.

1. Введение

Селективное зрительное внимание играет важную роль в жизнедеятельности животных и человека, позволяя выделять из большого объема зрительной информации ту часть, которая в данный момент наиболее значима и которая должна быть подвергнута более детальной обработке. Психологи выделяют несколько типов селективного внимания, которые соответствуют различным стратегиям формирования фокуса внимания. Пространственно-ориентированное внимание характеризуется в пространственных терминах, таких как "проектор" или "увеличительное стекло", которые могут быть направлены на оп-

* Работа поддержана Российским фондом фундаментальных исследований (проект 03-04-48482) и Президентской программой для ведущих научных школ (проект НШ 1872.2003.4).

ределенную часть зрительного поля [20, 48]. Объектно-ориентированное внимание привязано к определенным объектам независимо от того, в какой части пространства эти объекты находятся [19, 52].

Одной из экспериментальных парадигм в изучении объектно-ориентированного внимания является слежение за несколькими объектами (СНО). В традиционных экспериментах по СНО [49-52] испытуемому показывается экран дисплея, содержащий m простых по форме идентичных объектов (обычно используется до 10-12 объектов, таких как точки, круги, крестики и т.д.). В начале эксперимента некоторое подмножество этих объектов, содержащее от 1 до $m/2$ объектов (предполагается, что число m четное), коротко подсвечивается вспышками света. Подсвеченные таким образом объекты являются целевыми в том смысле, что именно за их перемещениями испытуемый должен следить на последующей стадии эксперимента. Остальные объекты считаются дистракторами. Дистракторы предназначены для повышения трудности слежения, поскольку они отвлекают внимание от целевых объектов.

После разделения объектов на целевые и дистракторы начинается следующий этап эксперимента, при котором объекты движутся независимо друг от друга в случайно выбранных и постоянно меняющихся направлениях. На этом этапе яркость всех объектов одинакова. Задача испытуемого состоит в том, чтобы следить за движениями целевых объектов при неподвижных зрачках (взгляд фиксируется на центр экрана). Время от времени один из объектов подсвечивается вспышкой света. В этот момент испытуемый должен нажать определенную клавишу в зависимости от того, идентифицирует ли он подсвеченный объект как целевой или дистрактор. Качество слежения характеризуется числом сделанных ошибок. В другой модификации эксперимента через фиксированное время (несколько секунд) перемещение объектов останавливается, и испытуемый должен с помощью компьютерной мышки указать все целевые объекты.

Проведенные в работе [50] эксперименты показали, что тренированные испытуемые способны следить за движением 2-5 целевых объектов с достаточно высокой надежностью. Хотя вероятность ошибки при идентификации растет с увеличением числа целей, даже в случае 5 целевых объектов вероятность правильной идентификации достигает 85%.

При интерпретации экспериментальных результатов было показано, что в данном случае имеет место параллельная обработка информации о целевых объектах, что соответствует гипотезе разделенного внимания. Последовательное циклическое переключение внимания по всем целевым объектам здесь вряд ли возможно, поскольку с учетом быстроты движения объектов перенос "проектора" внимания с одного объекта на другой должен был бы происходить с неправдоподобной с точки зрения нейрофизиологии скоростью.

Последующие работы по СНО [9, 40, 46, 53, 54, 65, 67] подтвердили, что первичная зрительная система способна "следить за движением до пяти зрительных объектов, причем осуществляется это без кодирования их визуальных характеристик" [49]. В работе [65] СНО исследовалось в случае, когда объекты могли пересекаться в пространстве при движении, так что их изображения накладывались одно на другое. Эксперименты проводились как при наличии "подсказки", характеризующей положение объектов по глубине, так и без такой

подсказки. Результаты экспериментов показывают, что "хотя задача слежения становится более трудной... тем не менее, она выполнима даже в случае перемещения объектов по двумерной плоскости". При наличии "подсказки" надежность СНО приближается к ситуации с непересекающимися объектами.

В настоящее время феномен внимания является популярным объектом нейросетевого моделирования. Разрабатываемые модели можно подразделить на две категории. Коннекционистские модели [25, 28, 29, 42, 47, 63], основанные на так называемой процедуре "победитель получает все" (winner-take-all), функционируют путем подходящей модификации весов связей в иерархической нейронной сети. Модели такого типа мало подходят для работы с движущимися объектами, поскольку для каждого нового положения объектов веса связей приходится пересчитывать заново.

Другую разновидность моделей внимания представляют осцилляторные нейронные сети [10, 11, 13, 32, 33, 66] (см. также обзорную статью [2]). Такие сети более подходят для моделирования объектно-ориентированного внимания, так как они оперируют в фазово-частотном пространстве. Благодаря этому достигается инвариантность фокусировки внимания по отношению к расположению объектов в физическом пространстве зрительного поля.

Модель СНО, представленная в данной работе, основана на разработанной нами ранее модели внимания с центральным осциллятором (МВЦО) [10, 11, 33-35]. Каждый элемент МВЦО представляет собой осциллятор, динамика которого описывается тремя переменными: фазой колебаний, амплитудой колебаний и собственной частотой осциллятора. Взаимодействие между осцилляторами реализуется в терминах фазовой синхронизации, резонанса и адаптации собственной частоты осциллятора.

МВЦО представляет собой однослойную сеть локально связанных осцилляторов, так называемых *периферических осцилляторов* (ПО), работа которых координируется *центральным осциллятором* (ЦО) [3, 17]. ЦО связан со всеми ПО прямыми и обратными связями. ПО представляют колонки кортикальных нейронов в первичных зонах зрительной коры (поля 17-18), реагирующие на определенные локальные признаки на изображении. Для простоты в качестве таких признаков используется контраст между интенсивностью заданного пикселя и интенсивностью фона. ЦО играет роль центрального управляющего элемента системы внимания (central executive) [5, 14]. В МВЦО изолированные объекты представляются в сети синхронными ансамблями ПО, а фокус внимания формируется теми ПО, которые работают синхронно с ЦО.

В психологической литературе центральный элемент представляет собой систему, осуществляющую управление рабочей памятью с помощью внимания [5-7, 56]. Вопрос о локализации этой системы в мозге до сих пор является предметом дискуссии. До недавнего времени считалось, что функцию центрального элемента выполняет определенная локальная область в префронтальной коре [17, 41]. Однако новые исследования дают основание предположить, что центральный элемент, по всей видимости, реализуется распределенной нейронной сетью, включающей латеральную префронтальную, орбитофронтальную и медиальную префронтальную зоны коры [4, 8]. Данные, полученные с помощью функциональной магнитной томографии и позитронно-эмиссионной томогра-

фии, показывают, что "различные управляющие функции не только связаны с определенными зонами фронтальной коры, но и зависят от задних (в основном теменных) областей" [12]. Помимо областей новой коры, также и гиппокамп может играть важную роль в реализации функций центрального элемента, поскольку: гиппокамп а) занимает финальное положение в пирамиде конвергентных зон коры [16], б) координирует обработку информации практически во всех областях неокортекса [27], в) координирует работу системы внимания [18, 26, 64].

С технической точки зрения наличие в системе центрального элемента, имеющего связи со всеми остальными элементами сети, позволяет сократить число связей в системе до величины порядка n , где n – число элементов в системе, в то время как для полностью связанной сети потребовалось бы число связей порядка n^2 . Это свойство делает системы с центральным элементом биологически реализуемыми даже в случае больших значений n .

В работах [11, 34] было показано, что с использованием принципов синхронизации и резонанса МВЦО может решать задачи последовательного выбора изолированных объектов в фокус внимания. В работе [1] МВЦО была использована для слежения за одним целевым объектом, движущимся среди объектов-дистракторов.

Идея одновременного слежения за m целевыми объектами состоит в том, чтобы объединить в единую сеть m копий МВЦО, каждая из которых должна следить за движениями одной цели. Реализация этой идеи упирается в решение следующих проблем. Во-первых, нужно предотвратить возникновение ситуации, когда один и тот же целевой объект одновременно попадает в фокус внимания нескольких МВЦО. Во-вторых, модель должна сохранять работоспособность в случае, когда объекты пересекаются в ходе своих перемещений. Поскольку объекты идентичны и движутся в случайных и постоянно изменяющихся направлениях, не существует безошибочной стратегии идентификации целевого объекта после того, как существенная его часть оказалась закрыта дистрактором. Разумной стратегией в данном случае является включение в фокус внимания любого из двух только что разошедшихся объектов, один из которых был целевым, а другой дистрактором. В случае существенного пересечения двух целевых объектов идентификация каждого из них также оказывается невозможной. Единственное, что может сделать в этом случае система внимания, это продолжать следить за обоими такими объектами после того, как они разойдутся. При этом возможно, что две МВЦО обменяются объектами, которые были у них в фокусе внимания. При следовании такой стратегии общее число объектов в фокусе внимания всегда остается равным m .

В разд. 2 будет описано, как согласовать работу нескольких МВЦО при слежении за набором из двух и более целевых объектов. В разд. 3, будут представлены результаты имитационных вычислений. Раздел 4 содержит обсуждение результатов.

2. Архитектура и принципы функционирования модели

Поскольку МВЦО является основным функциональным модулем модели СНО, опишем сначала архитектуру и принципы функционирования МВЦО. Архитектура МВЦО показана на рис. 1.

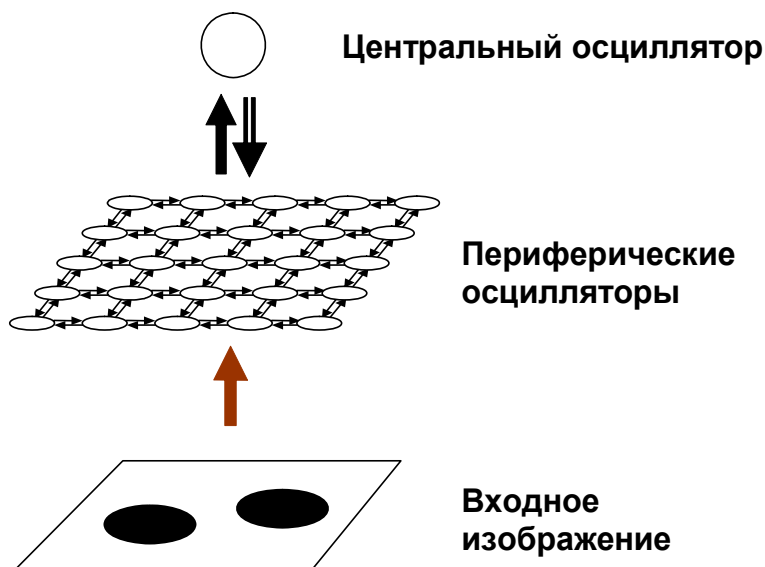


Рис. 1. Архитектура МВЦО. Коричневая стрелка показывает присвоение собственных частот периферическим осцилляторам. Черные стрелки показывают синхронизирующие связи, используемые (а) для связывания осцилляторов, кодирующих один объект, в синхронный ансамбль, (б) для синхронизации ансамбля ПО с ЦО. Полая стрелка показывает десинхронизирующие связи, используемые для предотвращения одновременной синхронизации ЦО с несколькими ансамблями ПО.

Входной информацией для сети является многоградационное монохроматическое изображение на плоской решетке пикселей. Решетка имеет те же размеры, что и сеть из периферических осцилляторов, так что каждый ПО получает внешний сигнал от пикселя, координаты которого идентичны координатам ПО. Для кодирования изображения в сети осцилляторов используется традиционный подход, согласно которому внешний сигнал задает собственную частоту осциллятора [38, 45]. Предполагается, что внешний сигнал формируется в латеральном генеральном ядре и зависит от контраста между интенсивностями пикселя и фона. В дальнейшем будут рассматриваться темные объекты на белом фоне, так что интенсивность фона всегда будет выше, чем интенсивность пикселей объектов. Величина собственной частоты осциллятора задается формулой

$$\omega_i = \lambda(B - I_i), \quad (0 \leq I_i \leq B),$$

где ω_i – собственная частота i -го ПО, I_i – интенсивность i -го пикселя, B – интенсивность фона, λ – шкалирующий параметр. Таким образом, большей величине контраста соответствует большее значение собственной частоты осциллятора.

ПО, соответствующие пикселям объектов, называются *активными*. Динамика этих осцилляторов задается уравнениями (1.1)–(1.4) из Приложения 1. ПО, соответствующие пикселям фона называются *молчащими*. Молчащие осцилляторы не принимают участия в динамике системы.

Каждый ПО имеет синхронизирующие локальные связи со своими ближайшими соседями. Эти связи используются для объединения в синхронный ансамбль периферических осцилляторов, представляющих изолированный объект. Такой способ представления объектов в нейронной сети соответствует синхронизационной гипотезе о связывании признаков объектов в единое целое [57].

ЦО имеет десинхронизирующие прямые и синхронизирующие обратные связи на каждый ПО. Синхронизирующие связи используются для синхронизации ЦО с ансамблем ПО. Десинхронизирующие связи используются для распределения ансамблей ПО в фазово-частотном пространстве так, чтобы разные ансамбли ПО работали несинхронно. Это препятствует одновременной синхронизации с ЦО нескольких ансамблей ПО.

Синхронизирующие и десинхронизирующие взаимодействия между ПО и ЦО порождают конкуренцию между различными ансамблями ПО за синхронизацию с ЦО. Только один ансамбль ПО может "победить" в этой борьбе, поэтому в каждый момент времени только один объект может быть включен в фокус внимания (исключение составляют лишь короткие периоды переходной динамики).

Синхронная работа ЦО и ансамбля ПО имеет следствием резонанс: амплитуды колебаний осцилляторов в ансамбле быстро возрастают до высокого уровня. Амплитуды остальных осцилляторов (не работающих синхронно с ЦО) подавляются до низкого уровня. Высокая амплитуда колебаний ПО является критерием того, что данный ПО включен в фокус внимания. Амплитуда колебаний ЦО является в случае МВЦО постоянной.

Поскольку ЦО может войти в синхронизацию только с теми ПО, чьи собственные частоты находятся в некотором диапазоне относительно собственной частоты ЦО, для настройки ЦО на частоту определенного ансамбля ПО необходимо, чтобы собственная частота ЦО совпадала с собственной частотой данного ансамбля ПО. В связи с этим в МВЦО собственная частота ЦО сделана адаптирующейся к текущей частоте ЦО. Благодаря этому ЦО как бы "ищет" в фазово-частотном пространстве тот ансамбль ПО, который является наиболее подходящим кандидатом на синхронизацию с ЦО.

Примеры численных экспериментов с МВЦО по выбору объектов в фокус внимания и по слежению за движением объекта можно найти в работах [1, 11, 34].

Перейдем теперь к описанию модели СНО. Архитектура соответствующей осцилляторной сети приведена на рис. 2.

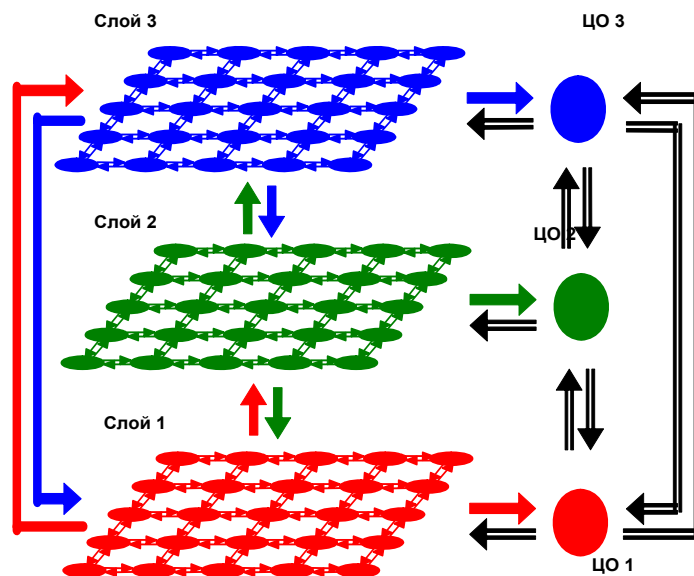


Рис. 2. Архитектура модели СНО. Слои сети раскрашены в разные цвета. Каждый слой предназначен для слежения за одним целевым объектом. Сплошные стрелки показывают синхронизирующие связи, полые стрелки показывают десинхронизирующие связи.

Показанная на этом рисунке сеть содержит 3 слоя (модуля) МВЦО, что соответствует случаю слежения за тремя целевыми объектами (в общем случае число слоев равно числу целевых объектов). Эти модули рассматриваются в модели как подсистемы системы внимания. Каждая подсистема предназначена для слежения за своим целевым объектом. В дальнейшем будет удобно использовать названия цветов (красный, синий, зеленый) для указания на конкретную подсистему внимания. Уравнения динамики модели приведены в Приложении 2 для общего случая m целевых объектов.

Периферические осцилляторы, имеющие одинаковые координаты на плоскости, но принадлежащие разным слоям, образуют колонку. ПО в колонке связаны сильными синхронизирующими связями типа "все на всех". Все ПО в колонке получают одинаковый внешний сигнал от соответствующего пикселя в зрительном поле. Как и в случае МВЦО, внешний сигнал определяется контрастом между интенсивностями в пикселе и на фоне. Этот сигнал задает собственные частоты ПО в колонке. В силу сформулированных условий собственные частоты в колонке идентичны, что способствует быстрой синхронизации осцилляторов в колонке.

Синхронизирующие локальные связи между ПО внутри слоя ограничены ближайшими соседями. Эти связи сделаны достаточно сильными, чтобы быстро синхронизировать все ПО, соответствующие изолированному объекту. Таким образом, наличие объекта в зрительном поле приводит к формированию синхронного ансамбля ПО, в который входят все колонки, активированные сигналами от данного объекта.

ЦО, принадлежащие различным слоям, соединены десинхронизирующими связями. Это сделано для того, чтобы предотвратить одновременную синхронизацию нескольких ЦО с одним и тем же ансамблем ПО при распределении целевых объектов между подсистемами внимания. Вследствие такой десинхронизации изолированные целевые объекты представляются в сети асинхронной активностью различных ансамблей ПО.

Как и в случае одного целевого объекта, если ансамбль ПО в k -ом слое работает синхронно с расположенным в этом слое ЦО, то амплитуды колебаний этих ПО быстро возрастают. Если амплитуда колебаний превышает заданный порог R , то считается, что осциллятор находится в резонансном состоянии. Если все ПО в k -ом слое, соответствующие пикселям некоторого объекта, находятся в резонансном состоянии, это интерпретируется как факт включения этого объекта в фокус внимания k -ой подсистемы внимания.

Опишем, что происходит с уже сформированным фокусом внимания при движении объектов. Если объекты движутся достаточно медленно и не пересекаются друг с другом, то фокус внимания обладает высокой устойчивостью. Дело в том, что резонансные ПО оказывают намного более сильное влияние на ЦО в своем слое, чем остальные ПО. Это препятствует переключению внимания на какой-либо ансамбль нерезонансных ПО. Ситуация изменяется, если скорость движения объектов становится достаточно большой по сравнению со скоростью протекания процессов синхронизации и возникновения резонанса. В этом случае возникают спонтанные переключения фокуса внимания с одного объекта на другой. Это приводит к ошибкам в идентификации целевых объектов и дистракторов.

Рассмотрим, что произойдет с фокусом внимания, если при движении имело место наложение двух объектов. Если оба объекта были вне фокуса внимания, то при их наложении фокус внимания не изменится. Временно образовавшийся дистрактор сложной формы так же не будет иметь возможность существенно влиять на фокус внимания, как и отдельные объекты-дистракторы.

Предположим, что произошло наложение двух объектов A и B , первый из которых входил в фокус внимания, а второй нет. В этом случае внимание распространится на весь составной объект, поскольку все ПО, соответствующие этому объекту будут включены в общий синхронный ансамбль ПО. Этот ансамбль будет работать синхронно с тем же ЦО, с которым был синхронизован ансамбль ПО, соответствующий объекту A до его пересечения с B .

После того как объекты A и B снова разойдутся и станут изолированными, соответствующие им ансамбли ПО возобновляют конкуренцию за то, чтобы быть включенными в фокус внимания. Поскольку связи от ЦО к ПО десинхронизирующие, только один из двух объектов имеет возможность выиграть в борьбе за синхронизацию с ЦО. Какой именно, зависит от случая – у системы нет информации, чтобы идентифицировать тот объект, который был в фокусе внимания до пересечения объектов. Если пересечение объектов A и B было достаточно значительным, то любой из них может быть выбран в фокус внимания с вероятностью 0.5, что, естественно, с той же вероятностью может повлечь возникновение ошибок при идентификации данных объектов.

Пусть произошло наложение двух объектов таких, что оба они были в фокусе внимания. В этом случае оба объекта останутся в фокусе внимания, несмотря на десинхронизирующие связи между центральными осцилляторами. Это достигается за счет того, что десинхронизация слаба по сравнению с синхронизирующим воздействием, получаемым центральными осцилляторами от общего ансамбля синхронно работающих ПО.

После того как объекты разошлись, десинхронизирующие связи между ЦО снова начинают играть решающую роль в динамике активности этих осцилляторов. Вследствие этого каждый объект "стремится" остаться в фокусе внимания ровно одного из модулей МВЦО. Конечно, вполне возможна такая ситуация, что объекты A и B , которые до пересечения друг с другом были в фокусе внимания, скажем, "красного" и "зеленого" модулей, поменяют следящие за ними подсистемы: после того как объекты A и B разойдутся, объект A окажется в фокусе внимания "зеленого" модуля, а объект B – "красного". Произойдет ли такая смена, зависит от того, насколько сильно пересеклись объекты A и B и сколько долго длилось это пересечение.

На самом деле, взаимодействие синхронизирующих и десинхронизирующих влияний на ЦО, реализованное в модели СНО, несколько более сложно, чем это было только что описано. Как показали вычислительные эксперименты, константная сила десинхронизирующего взаимодействия между центральными осцилляторами не может гарантировать нужной динамики ЦО. Если величина этой десинхронизирующей связи выбрана слишком большой, то ЦО может потерять синхронизацию с ансамблем ПО в тот момент, когда произошло наложение двух объектов, которые были в фокусе внимания. В результате только один ЦО сохранит синхронизацию с образовавшимся сложным объектом. Если же величина десинхронизирующей связи между ЦО выбрана слишком слабой, то после разделения сложного объекта оба ЦО могут оказаться синхронизованными с одним из объектов, а второй объект окажется вне фокуса внимания. Оказалось, что не удается подобрать константную величину связи между ЦО, которая бы позволила достаточно надежно избегать тех или иных из указанных ошибок.

С этой трудностью удалось справиться, предположив, что десинхронизирующее взаимодействие между ЦО усиливается после формирования сложного объекта из двух объектов, входивших в фокус внимания. В качестве механизма реализации этой гипотезы был использован резонанс, но в данном случае он был применен к центральному осцилляторам. Таким образом, в отличие от отдельной МВЦО, в модели СНО амплитуды колебаний ЦО не постоянны.

Резонансное возрастание амплитуды ЦО происходит, когда два объекта, находившихся в фокусе внимания, пересекаются друг с другом (см. уравнение (2.5) в Приложении 2). Благодаря последнему члену в уравнении (2.1) резонанс ЦО приводит к возрастанию десинхронизации между ЦО, в фокусе внимания которых находились указанные объекты. Поэтому в момент, когда объекты расходятся, десинхронизация между ЦО оказывается достаточно большой, чтобы предотвратить ситуацию, когда оба ЦО "следят" за одним объектом, в то время как второй объект остается вне фокуса внимания.

Для воспроизведения экспериментальных данных модель СНО помимо слежения за объектами должна быть способна выбрать в фокус внимания за-

данный набор целевых объектов. В экспериментальных условиях целевые объекты указываются испытуемому с помощью коротких вспышек света. Это делает целевые объекты заметными среди всего множества предъявленных объектов. В модели СНО для выбора объектов представление о заметности объектов формализуется в терминах силы действия периферических осцилляторов на центральные осцилляторы. Предполагается, что на этапе выбора целевых объектов те ПО, которые соответствуют пикселям, принадлежащим подсвечивавшимся объектам, оказывают на ЦО более сильное синхронизирующее воздействие, чем остальные ПО. В результате ансамбли ПО, соответствующие подсвечиваемым объектам, получают преимущество в конкуренции за синхронизацию с ЦО.

Идея о том, что фокусировка внимания связана с формированием в мозгу некоторой карты заметности была предложена в работе [36] и в дальнейшем активно использовалась в нейросетевых моделях зрительного поиска [28-30, 47]. Карта заметности представляет собой двумерную таблицу, которая кодирует заметность объектов среди их окружения в зрительной сцене и определяет приоритет включения объектов в фокус внимания. Для простоты в модели СНО карта заметности не вычисляется какой-либо нейронной схемой, а вводится формально как набор параметров s_i , соответствующих пикселям изображения и влияющих на силу воздействия ПО на ЦО, как это показано в уравнении (2.1) Приложения 2. Для различения подсвеченных и неподсвеченных объектов параметры s_i могут принимать одно из двух значений: большее значение $S_{flashed}$ соответствует пикселям подсвеченных объектов, меньшее значение $S_{non-flashed}$ соответствует пикселям неподсвеченных объектов. Для пикселей фона $s_i = 0$. Компьютерные вычисления показывают, что для того, чтобы с высокой степенью надежности обеспечить выбор в фокус внимания подсвеченных объектов, необходимо, чтобы величина $S_{flashed}$ была в несколько раз больше, чем $S_{non-flashed}$. В описываемых далее компьютерных экспериментах были использованы значения $S_{flashed} = 5$ и $S_{non-flashed} = 0,2$. После окончания подсветки целевых объектов и начала этапа движения все объекты имеют одинаковую заметность. Это отражается в модели заданием идентичных значений для всех s_i . В компьютерных экспериментах это значение полагалось равным 1.

3. Имитационные вычисления и сравнение с экспериментальными данными

Мы рассмотрим два случая функционирования модели СНО, соответствующие движениям объектов без пересечений и с пересечениями. В первом случае целью имитационного моделирования будет сравнить результаты функционирования модели с недавними экспериментальными данными по СНО, опубликованными в работе [46]. Во втором случае будет продемонстрировано, что в случае пересечения объектов фокус внимания в модели изменяется согласно процедурам, описанным в разделе 2. Тем самым, хотя число ошибок

слежения увеличивается, система слежения остается работоспособной в тех пределах, которые естественным образом вытекают из правил случайного движения объектов.

Авторы работы [46] экспериментировали с 12 объектами. Чтобы несколько ускорить вычисления мы ограничились 10 объектами, как это имело место в экспериментах [50]. Это не вносит существенных изменений в полученные результаты. В имитационных вычислениях в качестве объектов использовались черные квадраты размером 7×7 пикселей, расположенные на белом фоне в поле размером 30×60 пикселей. Рассматривались примеры с k целевыми объектами ($2 \leq k \leq 5$). В соответствии с этим использовались сети с k слоями (модулями МВЦО).

Параметры модели были подобраны таким образом, чтобы она более или менее адекватно отражала временные характеристики, типичные для психологических экспериментов по СНО. В качестве единицы времени был выбран интервал 100 мсек. Собственные частоты всех ПО были заданы равными $\omega_i = 5$ (5 колебаний за 100 мсек), что соответствует частоте 50 Гц. Таким образом, рабочей полосой частот модели является диапазон гамма-ритма. Амплитуды всех ПО в начальный момент устанавливались на значение 2, а в процессе работы системы варьировали в пределах (1, 11). Порог для резонансной амплитуды $R = 8.8$. Если амплитуда колебаний ПО превосходила R , считалось, что соответствующий пиксель включен в фокус внимания. Детальная информация о параметрах модели приведена в Приложении 2.

Время однократной работы модели (имеется в виду внутреннее время модели, а не длительность реальных вычислений на компьютере) будем называть *рабочим циклом*. Рабочий цикл составляет 7,2 сек и подразделяется на три фазы. Первая фаза (экспозиция) занимает 0,7 сек и используется для выделения целевых объектов. В это время объекты неподвижны. Единственное отличие целевых объектов от дистракторов задается с помощью карты заметности, которая формируется так, как это было описано в разделе 2. Ожидаемый результат работы модели при экспозиции состоит в том, что каждый целевой объект включается в фокус внимания одной из подсистем внимания. Ошибка может возникнуть в том случае, когда две подсистемы внимания фокусируются на одном целевом объекте или когда в фокус внимания одной из подсистем внимания включается объект-дистрактор. Компьютерные вычисления показали, что при выбранных параметрах модели такие ошибки весьма редки. Их вероятность не превосходит 0,005, поэтому практически они могут не приниматься во внимание.

Вторая фаза функционирования модели (слежение) занимает 6 сек. В течение этого времени объекты перемещаются в случайных направлениях. Скорость перемещения составляет 1 пиксель за 50 мсек, то есть каждые 50 мсек все квадраты подвергаются сдвигу на 1 пиксель в одном из четырех направлений: вверх, вниз, влево, вправо. Для каждого квадрата направление движения выбирается случайно и независимо с вероятностью 0,25. Для предотвращения столкновения квадратов вводится ограничение, согласно которому объекты должны отстоять друг от друга по меньшей мере на один пиксель. Сначала решается вопрос о том, будет ли объект двигаться в горизонтальном или вертикальном

направлении. Этот выбор делается с вероятностью 0,5. Для горизонтального движения затем решается вопрос, будет ли оно сделано влево или вправо. Этот выбор также делается с вероятностью 0,5. Аналогично, если принято решение о вертикальном движении, то его направление вверх или вниз выбирается с вероятностью 0,5. Если возникает опасность столкновения с другим объектом, направление движения меняется на противоположное. Если и в этом случае есть опасность столкновения, то в данный момент объект остается на месте. Аналогичные правила действуют для того, чтобы предотвратить пересечение объектом границы поля.

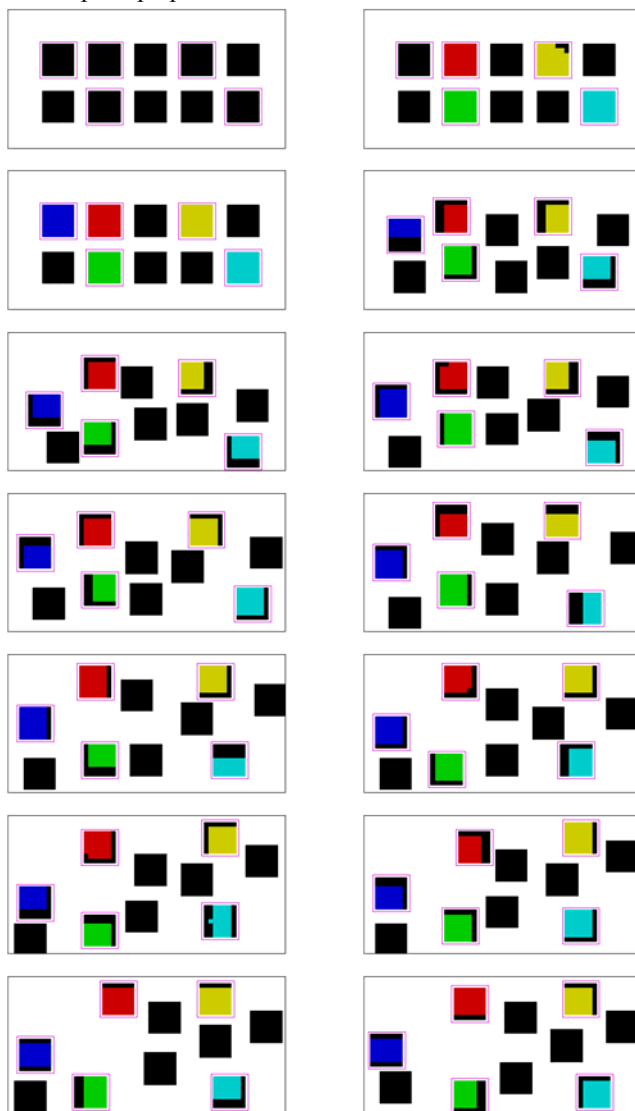
Третья фаза (тестирование) длится 0,5 сек. Во время этой фазы движение объектов прекращается. Эта фаза предназначена для того, чтобы дать системе время на принятие решения в ситуации, когда несколько объектов оказываются одновременно в фокусе внимания одной из подсистем внимания. Такая ситуация может время от времени возникать при движении объектов. Для стационарных объектов период времени 0.5 сек достаточен для того, чтобы подсистема внимания приняла решение, какой из этих объектов должен быть оставлен в фокусе внимания. Остальные объекты автоматически из фокуса внимания исключаются. Это достигается за счет уже упоминавшегося десинхронизирующего воздействия ЦО на ПО.

В финальный момент фазы тестирования происходит регистрация числа ошибок идентификации объектов как целевых и дистракторов. В соответствии с принятыми в данной модели принципами функционирования в этот момент в фокус внимания оказываются включенными ровно k квадратов, где k – число слоев в сети. Ошибки идентификации могут быть двух типов: целевой объект идентифицируется как дистрактор или дистрактор идентифицируется как целевой объект. В соответствии со стратегией, реализованной в модели, число ошибок всегда четно: если имеет место ошибка, когда целевой объект не попадает в фокус внимания, это неизбежно влечет включение в фокус внимания какого-то дистрактора.

На рис.3 показаны кадры из фильма, иллюстрирующего динамику формирования и переключения фокуса внимания при слежении за пятью целевыми объектами при наличии пяти дистракторов. Кадры упорядочены слева направо и сверху вниз. Временной интервал между кадрами 0,3 сек. Целевые объекты показаны с обведенной вокруг них сиреневой рамочкой. Объекты, находящиеся в текущий момент в фокусе внимания, покрашены в различные цвета. Цвет пикселя в каждом кадре зависит от состояния ПО в колонке, соответствующей этому пикселю. Если ПО в "красном" слое находится в резонансном состоянии, тогда пиксель окрашен в красный цвет. Аналогичный принцип используется для остальных цветов, кроме черного и белого. В черный цвет покрашены пиксели, находящиеся вне фокуса внимания. В белый цвет покрашены пиксели фона.

Левая верхняя рамка соответствует начальному моменту, когда в фокусе внимания еще нет объектов. Фокус внимания, включающий пять целевых объектов формируется на первом кадре второй строки. Второй кадр второй строки соответствует началу фазы слежения. Первоначально модель удерживает в фокусе внимания все изначально заданные целевые объекты, но в конце фазы

слежения происходит ошибочное переключение "зеленой" подсистемы внимания с целевого объекта на один из объектов дистракторов. Позднее эта подсистем внимания захватывает два объекта-дистрактора в фокус внимания, но на этапе тестирования (последний кадр в последней строке) в фокусе внимания этой подсистемы остается один дистрактор. В результате число ошибок идентификации в этом примере равно 2.



Нейросетевая модель слежения за несколькими объектами

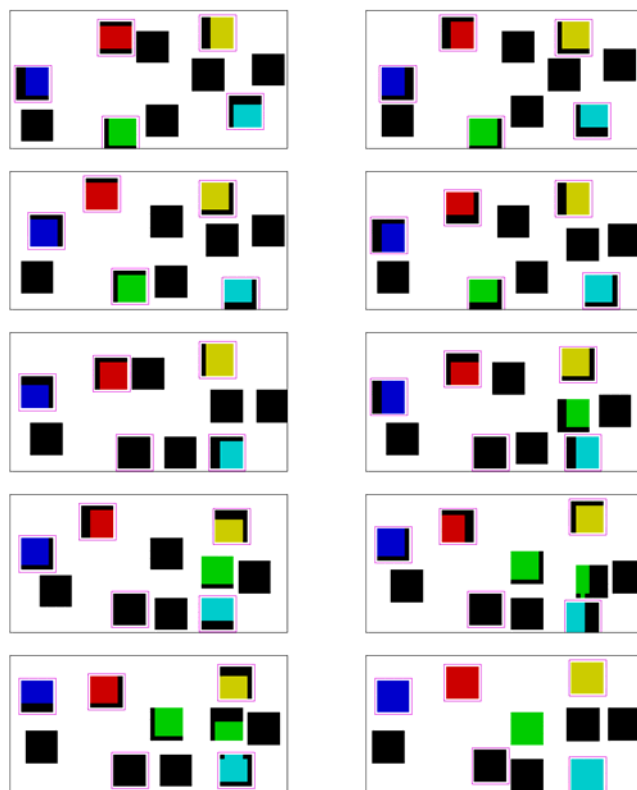


Рис. 3. Слежение за несколькими объектами при запрете на наложение объектов. Изображение содержит 5 целевых объектов и 5 дистракторов. Пиксели объектов, включенные в фокус внимания, окрашены в различные цвета, соответствующие различным подсистемам внимания. Пиксель имеет черный цвет, если ни один из осцилляторов в колонке не находится в резонансном состоянии. Пиксель имеет белый цвет, если он принадлежит фону.

Для оценки качества функционирования модели в терминах числа ошибок идентификации для каждого значения числа целевых объектов $k = 2, 3, 4, 5$ было проделано по 50 рабочих циклов модели. Результаты вычислений показаны в таблице 1.

Таблица 1. Результаты идентификации целевых объектов и дистракторов моделью СНО

	Число целевых объектов			
	2	3	4	5
Общее число ошибок	2	38	86	134
Среднее число ошибок на рабочий цикл	0.04	0.76	1.72	2.68
Стандартное отклонение	0.3	1.1	1.5	1.6
Вероятность ошибки идентификации пробного объекта	0.006	0.09	0.179	0.268

Для проверки статистической значимости отличия средних значений числа ошибок в четырех группах, соответствующих четырем значениям k , был использован тест анализа дисперсии (ANOVA). Проверяемая гипотеза состояла в том, что все группы (каждая из которых содержит по 50 чисел) были получены из одного вероятностного распределения (или из разных распределений, но имеющих одно и то же среднее). Полученный результат: $F = 43,7$ и $p < 0,0001$. Это означает, что гипотеза должна быть отвергнута.

Дальнейший анализ с помощью парного Т-теста дал следующие результаты: $T_{23} = 4.6$, $T_{34} = 3.7$, $T_{45} = 3.1$. Из этого следует, что гипотезы о том, что среднее в группе k равно среднему в группе $(k-1)$ для $k = 3,4,5$ должны быть отвергнуты.

В экспериментах [46] результат СНО для человека оценивался с помощью пробного объекта, который подсвечивался в финальный момент каждого опыта и должен был быть идентифицирован испытуемым как целевой или дистрактор. Чтобы исключить возможность угадывания на основе учета соотношения между числом целевых объектов и дистракторов, пробный объект выбирался из набора целевых объектов или из набора дистракторов с вероятностью 0,5. В наших экспериментах пробные объекты не использовались, однако наши данные относительно числа ошибок в каждом рабочем цикле дают возможность вычислить вероятность ошибки при использовании пробного объекта.

Пусть s – число объектов ($s = 10$), k – число целевых объектов ($k = 2,3,4,5$), e – число целевых объектов, которые были ошибочно идентифицированы как дистракторы в одном рабочем цикле (следовательно, то же число дистракторов было ошибочно идентифицировано как целевые объекты). Тогда вероятность ошибки идентификации при использовании пробного объекта будет

$$P = 0.5 \left(\frac{e}{k} + \frac{e}{s-k} \right) = \frac{0.5se}{k(s-k)}.$$

Используя эту формулу, мы вычислили значения P для каждого рабочего цикла и затем усреднили эти значения по всем рабочим циклам. Результаты представлены в последней строке таблицы 1 и на рис. 4. Для сравнения на рис. 4 показаны также данные о качестве СНО для людей, приведенные в [46].

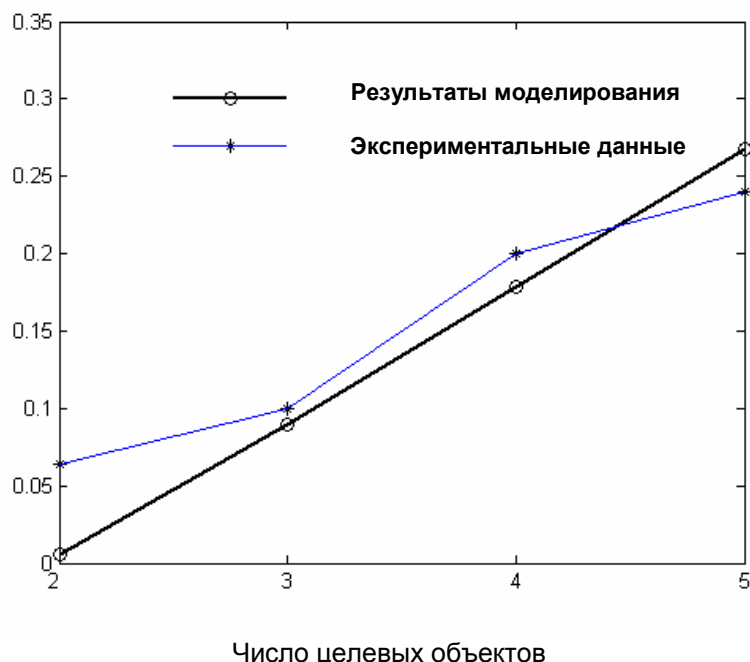


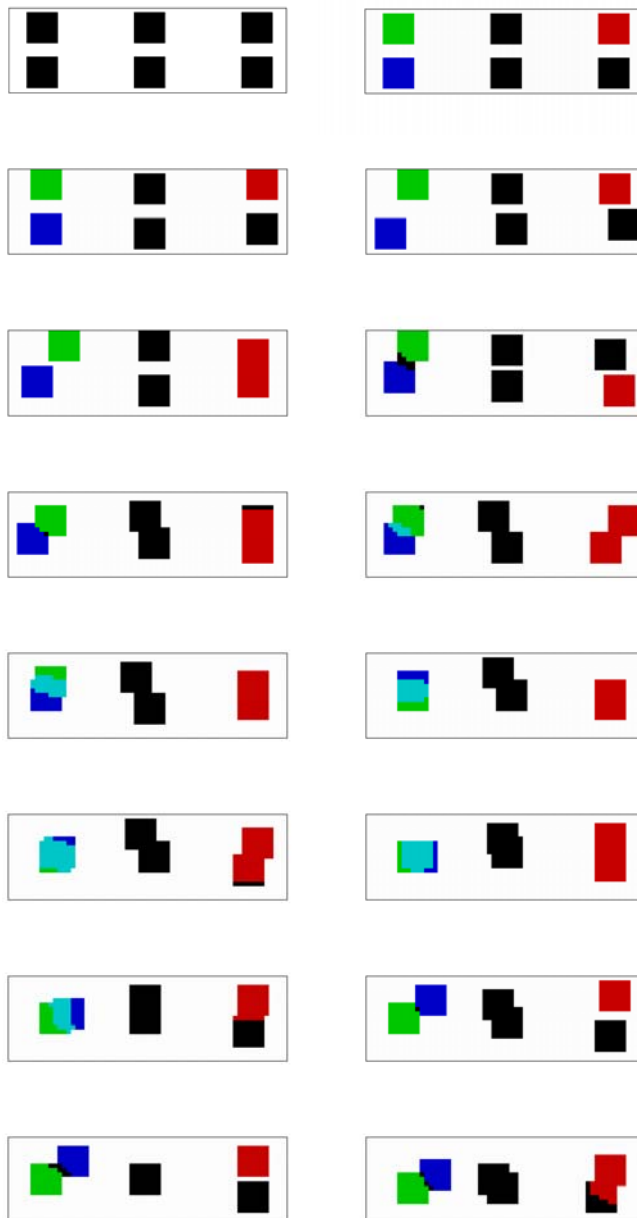
Рис. 4. Вероятность ошибки при тестировании пробного объекта в зависимости от числа целевых объектов. Экспериментальные данные взяты из работы [46].

Оба графика на рис. 4 имеют сходное поведение: вероятность ошибки возрастает с увеличением числа целевых объектов. Основное отличие между графиками имеет место для случая двух целевых объектов. В этом случае модель дает существенно меньшую вероятность ошибки, чем эксперимент.

Проиллюстрируем теперь функционирование модели в случае, когда допускается пересечение объектов при их движении. В качестве примера рассмотрим слежение за 6 объектами (3 целевых объекта и 3 дистрактора). Как и прежде, объектами являются черные квадраты размером 7 x 7 пикселей, расположенные на белом фоне. Размер поля 19 x 62 пикселей.

Рабочий цикл модели разделен на две фазы. Фаза экспозиции (1,2 сек) используется для указания целевых объектов и выбора их в фокус внимания. В этой фазе все квадраты представляют собой неподвижные изолированные объекты. Выбор целевых объектов осуществляется так же, как это было описано выше.

В фазе слежения, которая длится 9,6 сек, объекты движутся аналогично тому, как это было описано выше, но без ограничений на пересечение объектов. На рис. 5 показаны кадры из фильма, иллюстрирующего динамику формирования и переключения фокуса внимания.



Нейросетевая модель слежения за несколькими объектами

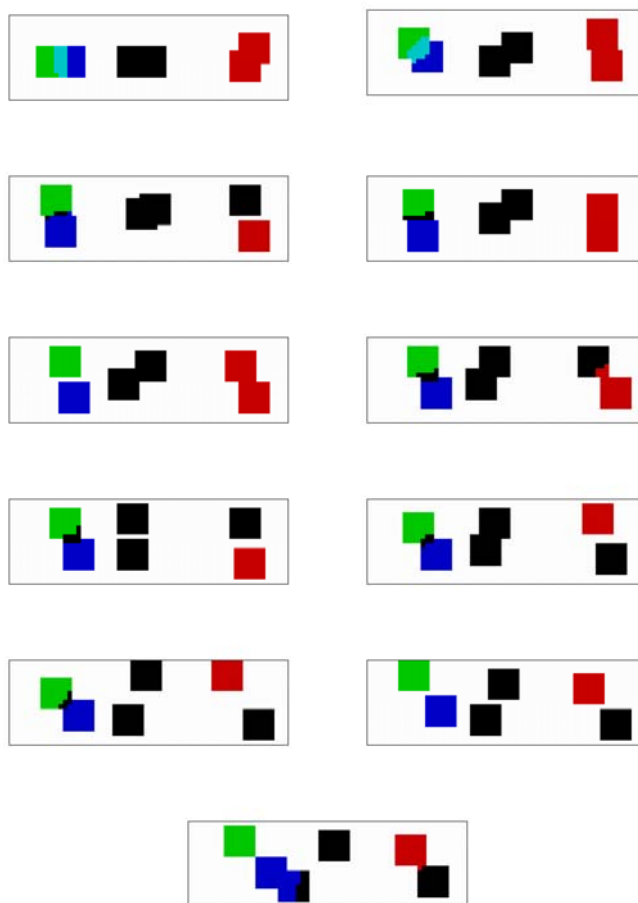


Рис. 5. Слежение за несколькими объектами при разрешении наложения объектов. Изображение содержит 3 целевых объекта и 3 дистрактора. Каждый пиксель имеет один из цветов: красный, синий, зеленый или смесь этих цветов. Пиксель имеет красный/зеленый/синий цвет, если представляющий этот пиксель ПО в "красном"/"зеленом"/"синем" слое находится в резонансном состоянии. Пиксель имеет голубой цвет, если оба осциллятора в "синем" и "зеленом" слоях находятся в резонансном состоянии. Пиксель имеет черный цвет, если ни один из осцилляторов в колонке не находится в резонансном состоянии. Пиксель имеет белый цвет, если он принадлежит фону.

Кадры упорядочены слева направо и сверху вниз. Временной интервал между кадрами 0,4 сек. Кадр в левом верхнем углу показывает начальное положение квадратов на изображении. На следующем кадре показан момент, когда формируется фокус внимания, содержащий 3 целевых объекта. Первый кадр во второй строке кадров соответствует началу фазы движения объектов. В даль-

нейшем это движение приводит к образованию различных комбинаций пересекающихся объектов.

Как и на рис. 3 цвета на рис. 5 это не цвета, в которые окрашены объекты (все объекты на входном изображении черные). Цвета используются для того, чтобы показать, какие объекты находятся в фокусе внимания той или иной подсистемы внимания. Цвет пикселя в каждом кадре зависит от состояния ПО в колонке, соответствующей этому пикселю. Если ПО в "красном" слое находится в резонансном состоянии, тогда пиксель окрашен в красный цвет. Аналогично для обозначения резонансных осцилляторов используются зеленый и синий цвета. Если в колонке несколько осцилляторов находятся в резонансном состоянии, то цвет пикселя формируется из смеси основных цветов. Например, пиксель имеет голубой цвет, если он одновременно включен в фокус внимания "зеленой" и "синей" подсистем внимания. Черный цвет используется для пикселей объектов вне фокуса внимания, белый цвет используется для пикселей фона. (Интенсивности зеленого, красного и синего подобраны так, чтобы их комбинация в одном пикселе давала серый цвет. На самом деле, на рис. 4 таких пикселей нет).

Рассмотрим, что происходит с фокусом внимания в процессе перемещения объектов.

Пара квадратов, расположенных в центре изображения, всегда находится вне фокуса внимания независимо от того, являются ли эти квадраты изолированными объектами или временно формируют сложный объект. ПО, представляющие пиксели этих объектов, всегда работают с низкой амплитудой.

Пара квадратов в правой части изображения показывает процесс переключения внимания в случае, когда сложный объект образуется в результате пересечения двух объектов, один из которых находится в фокусе внимания, а другой нет. Если время существования сложного объекта невелико, то после расхождения квадратов внимание остается на том же объекте, на котором оно было сфокусировано до наложения объектов. Если область пересечения объектов достаточно велика и время существования сложного объекта достаточно продолжительное, то это приводит к тому, что внимание распространяется на весь сложный объект. После того как квадраты снова разойдутся, любой из них может остаться в фокусе внимания, тогда как другой квадрат будет исключен из фокуса внимания.

Наконец, рассмотрим пару квадратов в левой части изображения. Эти квадраты оба были указаны как целевые объекты и вследствие этого они были выбраны в фокус внимания двух подсистем внимания "зеленой" и "синей". До тех пор пока эти квадраты не пересекаются, их движение устойчиво отслеживается этими подсистемами. Ситуация меняется после того, как возникает сложный объект с достаточно большой областью пересечения этих квадратов. Эта область на рисунке окрашена в голубой цвет, поскольку ее пиксели одновременно находятся в фокусе внимания обеих подсистем, "зеленой" и "синей". В процессе перемещения объектов размер этой голубой области меняется, она может увеличиваться или уменьшаться, исчезать и снова появляться. Существенно лишь то, что после того как квадраты разойдутся, движение каждого из них будет вновь отслеживаться одной из подсистем внимания. Обмен целевых

объектов между подсистемами внимания может произойти тогда, когда область пересечения квадратов в сложном объекте достаточно велика и не позволяет надежно идентифицировать каждый из квадратов после того, как они разойдутся.

Количественное исследование качества функционирования модели СНО в случае пересекающихся объектов является предметом дальнейших исследований. Предварительный опыт имитационных вычислений показал, что вероятность ошибок существенно зависит от скорости перемещения объектов. Для невысоких скоростей (именно такие скорости использовались в приведенном выше примере), вероятность ошибок (за исключением тех, которые неизбежны при пересечении целевых объектов и дистракторов) достаточно мала.

4. Обсуждение

Представленная в данной работе модель внимания основана на трех основных идеях: Во-первых, в качестве ключевых механизмов формирования фокуса внимания нами используются колебания и их синхронизация. Экспериментальное подтверждение того, что осцилляторная активность и фазовая синхронизация имеют прямое отношение к феномену селективного зрительного внимания, было получено в работах по регистрации электроэнцефалограмм [26], магнитных энцефалограмм [24, 58], и локальных полевых потенциалов [21, 22, 23, 39, 44, 60]. В частности было показано, что модуляция синхронизации нейронной активности определяет исход при решении поведенческой задачи, требующей привлечения внимания [24, 61].

Во-вторых, в модели используется гипотеза о существовании специальной нейронной системы, центрального управляющего элемента координирующего фокусировку внимания. Теоретические и экспериментальные соображения в пользу этой гипотезы приведены в [5-7, 14, 56]. В терминах модели зрительное внимание обычно характеризуется как следствие синхронизации активности центрального элемента и связанного ансамбля нейронов в первичных зонах зрительной коры. СНО дает пример более сложной ситуации, когда внимание распределено между несколькими изолированными объектами. Мы предполагаем, что в этом случае центральный элемент как бы разбивается на несколько подсистем, работающих асинхронно. Синхронные же колебания используются в качестве "метки", позволяющей различать объекты, одновременно включенные в фокус внимания.

В третьих, в модели используется резонанс. Его назначение – формализовать гипотезу об увеличении уровня нейронной активности в областях мозга, включенных в фокус внимания, и подавить до низкого уровня активность, возникающую в ответ на стимулы вне фокуса внимания.

Связь между резонансом в модели и модуляцией кортикальной активности при внимании следует прояснить более подробно. Нейрофизиологические [43, 51] и томографические исследования [31, 55, 59] показали, что связанная с вниманием модуляция может проявляться уже на уровне первичных зон зрительной коры, однако проявление эффектов внимания увеличивается при движении вверх по иерархии кортикальных структур [62]. Более того, в работе [15]

функциональная магнитная томография была применена для исследования уровня нейронной активности, возникающей в экспериментах с фокусировкой внимания на движущихся объектах. Авторы не обнаружили какого-либо увеличения активности в первичной зрительной коре, но в теменных и фронтальных зонах коры активность по меньшей мере удваивалась. Как привести эти данные в соответствие с динамикой модели СНО?

В модели были использованы непосредственные связи от ПО к ЦО. Это, конечно, является сильным упрощением реальной ситуации. На самом деле, путь от первичных зон зрительной коры до областей, занятых центральным управляющим элементом проходит через многие промежуточные кортикальные структуры. К моменту, когда поток информации достигает верхних уровней коры, разница в активности между тем, что находится или не находится в фокусе внимания, становится вполне очевидной. И все же эта разница не так велика, как та, что имеет место в модели для резонансных и нерезонансных колебаний. Поэтому не следует воспринимать амплитуду колебаний в модели как реальное отражение наблюдаемой в эксперименте кортикальной активности. Амплитуду колебаний в модели следует скорее рассматривать как формальную характеристику, положительно коррелирующую с активностью в коре и определяющую силу взаимодействия между кортикальными осцилляторами и центральным управляющим элементом.

Теоретическая концепция, используемая для объяснения СНО, основана на идее индексации объектов, подлежащих слежению, на уровне предвнимания [49, 50]. Предполагается, что индексация происходит независимо и параллельно в нескольких участках зрительного поля. В отличие от этой теории в нашей модели индексация есть результат одновременной работы предвнимания и внимания. Предвнимание ответственно за то, чтобы присвоить каждому объекту метку в виде определенной колебательной активности. Вся информация об объекте кодируется в форме синхронных (синфазных) колебаний. Хотя такая метка в определенных пределах меняется во времени, она позволяет достаточно надежно отличать даже идентичные по своим признакам объекты. Различение объектов и выделение из них определенного подмножества, предназначенного для каких-то специальных процедур (в данном случае слежения) осуществляется системой внимания. В случае, когда в фокусе внимания должно находиться несколько объектов, система внимания функционирует как несколько взаимодействующих подсистем. Номер подсистемы, в фокусе внимания которой находится некоторый объект, может рассматриваться как индекс этого объекта.

Рассматриваемая модель СНО имеет жесткую архитектуру связей и априорно заданные и не изменяющиеся функции взаимодействия и параметры. Реальная биологическая система, реализующая СНО, несомненно, должна быть намного более гибкой. В частности, модель СНО с фиксированным числом модулей МВЦО не может адаптироваться к слежению за произвольно заданным числом целевых объектов. Конечно, можно предположить, что число активных ЦО может задаваться с помощью внутренней психологической установки и всегда соответствует числу целевых объектов. Нам, однако, представляется, что более правдоподобным решением является гибкая, зависящая от решаемой задачи, архитектура взаимодействий при формировании системы центральных

осцилляторов. Если необходимо следить за одним объектом, все центральные осцилляторы объединяются в один ансамбль с помощью синхронных взаимодействий. Если требуется следить за двумя объектами, то формируется две группы центральных осцилляторов, с синхронизирующими взаимодействиями между осцилляторами в группе и десинхронизирующими взаимодействиями между группами и т.д.

Зависимость вероятности ошибок в СНО от числа целевых объектов является причиной, по которой исследователи-экспериментаторы рассматривают СНО как пример параллельной обработки информации с ограниченными ресурсами [50]. Модель СНО показывает, чем может быть обусловлена эта ограниченность. Увеличение вероятности ошибок в модели связано с ограниченной емкостью фазового пространства, в котором одновременно вынуждено оперировать несколько ЦО. Увеличение числа ЦО приводит к тому, что им все труднее становится избегать синхронизации друг с другом, что приводит к непредсказуемым переключениям внимания на объекты-дистракторы. Пока число целевых объектов не превышает 5, вероятность таких переключений мала, если объекты стационарны, но эта вероятность существенно возрастает, когда объекты начинают двигаться с достаточно большой скоростью. Из-за перемещений объектов процессы синхронизации и резонанса не могут продвинуться достаточно далеко, что приводит к потере синхронизации ЦО с выбранным ранее в фокус внимания ансамблем ПО.

Демонстрируемое моделью в случае двух целевых объектов существенно лучшее качество слежения по сравнению с людьми, показывает, что в этом случае, вероятно, есть какие-то дополнительные факторы, которые делают данную задачу для людей достаточно сложной.

Хотя параметры модели были выбраны таким образом, чтобы временные соотношения приближались к тем, которые используются в экспериментах СНО, к этому факту не стоит относиться слишком серьезно. Модель слишком груба, чтобы в данном аспекте быть точной. Весьма небольшое число пикселей в объектах, ограничение взаимодействия ПО ближайшими соседями и многие другие особенности модели обусловлены необходимостью уложиться в разумное время при компьютерных вычислениях. На самом деле, модель довольно гибка в отношении воспроизведения различных временных показателей. Другой выбор параметров или даже просто изменение единицы времени позволяет увеличить или уменьшить длительность рабочего цикла. В связи с этим, при сравнении качества функционирования модели по сравнению с людьми мы использовали экспериментальные данные, усредненные по периодам времени 5, 9 и 13 сек. Данные для 5 сек в экспериментах с людьми давали меньшие вероятности ошибок, чем те, что были получены в наших вычислениях, но поведение этих вероятностей при увеличении числа целевых объектов остается тем же самым.

Использованная в модели стратегия принятия решений также является сильно упрощенной. Модель всегда старается следить за фиксированным и заданным заранее числом объектов. Эксперименты показывают, что стратегия людей более разумна [46, 50]. Если испытуемый чувствует, что правильная идентификация объекта в процессе слежения становится сомнительной, он

склонен прекратить слежение этого объекта и сосредоточить внимание на слежении меньшего числа объектов. Это может служить объяснением тому, что качество слежения для человека падает не резко, а постепенно, даже когда число целевых объектов превосходит 5. Этот же факт может быть причиной того, что качество слежения для человека при переходе от четырех целевых объектов к пяти падает меньше, чем в модели СНО.

Известно, что в экспериментах СНО качество слежения может быть улучшено, если испытуемый группирует целевые объекты в виртуальный многоугольник, а затем следит за деформациями этого многоугольника. [67]. Внешне слежение в этом случае выглядит весьма сходно, однако при его реализации могут быть задействованы более сложные когнитивные механизмы, которые, по меньшей мере, должны включать наличие у испытуемого представления о геометрических фигурах. Границы виртуального многоугольника являются воображаемыми линиями, которые формируются в результате внутреннего усилия. После этого многоугольник может рассматриваться как связный объект, фокусировка внимания на котором не требует разделенного внимания и может быть выполнена в рамках МВЦО. Трудность слежения в этом случае обусловлена, по-видимому, другими причинами, чем те, что рассмотрены в данной работе.

При разработке модели СНО мы старались исключить из операций, совершаемых моделью, традиционные процедуры обработки изображений, такие как анализ формы объектов, тестирование связности, распознавание и т.д. Благодаря этому модель может с равным успехом работать с неидентичными объектами или объектами, форма которых меняется во времени. Последнее важно, например, если объекты движутся в трехмерном пространстве, когда проекция их на сетчатку постоянно меняется. Другой причиной нашего подхода к моделированию СНО является тот факт, что перечисленные виды обработки зрительной информации требуют большего времени и требуют участия более высоких кортикальных структур. Нашей задачей было исследовать, может ли СНО быть объяснено в терминах простой нейронной сети, в которой единственная роль, выполняемая высокой структурой, состоит в том, чтобы направлять внимание на нужные объекты.

Приложение 1. Математическое описание МВЦО

Осцилляторы, из которых строится сеть МВЦО, представляют собой обобщенные фазовые осцилляторы. Состояние такого осциллятора описывается тремя явно заданными переменными: фазой колебаний, амплитудой колебаний и собственной частотой осциллятора (осциллятор генерирует колебания с собственной частотой при отключении взаимодействия с другими осцилляторами сети).

Динамика МВЦО описывается следующими уравнениями

$$\frac{d\theta_0}{dt} = 2\pi\omega_0 + \frac{w_0}{n} \sum_{i=1}^n s_i a_i g(\theta_i - \theta_0), \quad (1.1)$$

Нейросетевая модель слежения за несколькими объектами

$$\frac{d\theta_i}{dt} = 2\pi\omega_i - a_0 w_1 h(\theta_0 - \theta_i) + w_2 \sum_{j \in N_i} a_j p(\theta_j - \theta_i) + \rho, \quad (1.2)$$

$$\frac{da_i}{dt} = \beta(-a_i + \gamma f(\theta_0 - \theta_i)), \quad (1.3)$$

$$\frac{d\omega_0}{dt} = -\alpha(2\pi\omega_0 - \frac{d\theta_0}{dt}). \quad (1.4)$$

В этих уравнениях θ_0 – фаза ЦО, θ_i ($i = 1, \dots, n$) – фазы ПО, $\frac{d\theta_0}{dt}$ и $\frac{d\theta_i}{dt}$ – текущие частоты осцилляторов, ω_0 – собственная частота ЦО, ω_i – собственные частоты ПО, a_0 – амплитуда колебаний ЦО (константа), a_i – амплитуды колебаний ПО, w_0, w_1, w_2 , – положительные константы, задающие силу взаимодействия между осцилляторами, s_i – параметр, задающий различие между активными и молчащими ПО, $s_i = 1$, если ПО активен, в противном случае $s_i = 0$, N_i – набор активных ПО в ближайшей окрестности осциллятора i , ρ – гауссовский шум со средним 0 и стандартным отклонением σ , функции g, h, p задают взаимодействие между осцилляторами, f – функция, задающая амплитуду колебаний периферических осцилляторов и их переход в резонансное состояние, α, β, γ – параметры (положительные константы). Значения ω_i задаются внешними сигналами на ПО; $\theta_0, \theta_i, \omega_0, a_i$ – внутренние переменные, задающие состояние сети.

Функции g, h, p 2π – периодические, нечетные, одномодальные в положительном полуинтервале периодичности, функция f 2π – периодическая, четная, положительная, одномодальная в интервале периодичности.

Приведем формулы для перечисленных функций, использовавшиеся при имитационных расчетах. На интервале $(0, \pi)$ эти функции имеют вид

$$g(x) = \begin{cases} 10x, & 0 \leq x < 0.1, \\ -4x + 1.4, & 0.1 \leq x < 0.2, \\ -0.1x + 0.62, & 0.2 \leq x \leq \pi, \end{cases}$$

$$h(x) = 4.8x \exp(-4.8x + 1),$$

$$p(x) = \sin x,$$

$$f(x) = S((\cos x)^+),$$

где

$$y^+ = \begin{cases} y, & y \geq 0, \\ 0, & y < 0, \end{cases}$$

$$S(z) = \frac{\exp((z - 0.9)/0.02)}{1 + \exp((z - 0.9)/0.02)} + 0.1.$$

На интервал $(-\pi, 0)$ функции g, h распространяются как нечетные, а функция f как четная.

Уравнения (1.1)–(1.2) – традиционные уравнения фазовой синхронизации, соответствующие архитектуре связей, приведенной на рис. 1. Эти уравнения описывают процессы синхронизации/десинхронизации в сети. Уравнение (1.1) задает динамику ЦО. Уравнение (1.2) задает динамику периферических осцилляторов. Шум ρ в (1.2) используется как дополнительный источник десинхронизации между синхронными ансамблями ПО. Шум помогает рандомизировать расположение различных ансамблей ПО в фазово-частотном пространстве, благодаря чему ЦО способен эти ансамбли различить.

Уравнение (1.3) описывает динамику амплитуд ПО. Интервал вариации функции $f - (0, 1; 1, 1)$. Поэтому амплитуда колебаний ПО возрастает до максимального значения $a_{\max} = 1,1\gamma$, если ПО работает синхронно (синфазно) с ЦО, и падает до низких значений $a_{\min} = 0,1\gamma$, если фаза ПО существенно отличается от фазы ЦО. Считается, что осциллятор находится в резонансном состоянии, если амплитуда колебаний превосходит порог $R = 0,8a_{\max}$. Параметр β задает скорость изменения амплитуды.

Уравнение (1.4) описывает адаптацию собственной частоты ЦО. В соответствии с этим уравнением значение $2\pi\omega_0$ стремится к текущей частоте ЦО. Благодаря этой адаптации ЦО способен "искать" в фазово-частотном пространстве ансамбль ПО, наиболее подходящий для синхронизации

Приложение 2. Математическое описание модели СНО

Уравнения для динамики модели СНО являются модификацией уравнений (1.1)–(1.4) применительно к архитектуре связей, представленной на рис. 2,

$$\frac{d\theta_0^k}{dt} = 2\pi\omega_0^k + \frac{w_0}{n_{res}} \sum_{i=1}^n s_i a_i^k g(\theta_j^k - \theta_0^k) - w_3 \sum_{l=1}^m a_0^l h(\theta_0^l - \theta_0^k), \quad (2.1)$$

$$\frac{d\theta_i^k}{dt} = 2\pi\omega_i^k - a_0^k w_1 h(\theta_0^k - \theta_i^k) + w_2 \sum_{j \in N_i} a_j^k p(\theta_j^k - \theta_i^k) + \frac{w_4}{m} \sum_{l=1}^m a_l^l p(\theta_i^l - \theta_i^k) + \rho, \quad (2.2)$$

$$\frac{da_i^k}{dt} = \beta(-a_i^k + \gamma f(\theta_0^k - \theta_i^k) + \zeta), \quad (2.3)$$

$$\frac{d\omega_0^k}{dt} = -\alpha(2\pi\omega_0^k - \frac{d\theta_0^k}{dt}). \quad (2.4)$$

В этих уравнениях верхний и нижний индексы нумеруют, соответственно, слои и осцилляторы в слое; $k = 1, \dots, m$, где m – число слоев. Нормирующий

параметр n_{res} равен числу резонансных осцилляторов, но не меньше, чем 49 (число пикселей в одном объекте). Последний член в (2.1) описывает взаимодействие между центральными осцилляторами (знак минус перед этим членом указывает на то, что взаимодействие десинхронизирующее). Член, предшествующий шуму ρ в (2.2), описывает взаимодействие ПО внутри колонки. Параметры w_3, w_4 – положительные константы.

Параметры s_i в уравнении (2.1) формируют карту заметности, $s_i > 0$ для пикселей объектов, $s_i = 0$ для пикселей фона. В фазе экспозиции значения $s_i = 5$ для пикселей целевых объектов, $s_i = 0,2$ для пикселей дистракторов. В фазе слежения и тестирования все $s_i = 1$.

Уравнения (2.3)-(2.4) являются обобщением уравнений (1.3)-(1.4) для случая многослойной сети.

Амплитуды ЦО меняются в соответствии с уравнением, сходным с уравнением (2.3),

$$\frac{da_0^k}{dt} = \beta(-a_0^k + \gamma_1 r(\sum_{l=1, l \neq k}^m f(\theta_0^l - \theta_0^k))) + \zeta, \quad (2.5)$$

где ζ - положительная константа,

$$r(x) = \begin{cases} x, & x \leq 1, \\ 1, & x > 1. \end{cases}$$

Функция r необходима для стандартизации интервала вариации амплитуд ЦО, который в данном случае равен $(0, 1\gamma_1 + \zeta, \gamma_1 + \zeta)$.

Параметры уравнений (2.1)-(2.5) и начальные значения переменных приведены в таблицах 2 и 3.

Таблица 2. Параметры, использовавшиеся при имитационном моделировании.

Параметры	Значения	Параметры	Значения
w_0	5	w_2	40
w_3	-12 (при экс- позиции) -7 (при сле- жении и тес- тировании)	w_4	40
α	2	β	1
γ	10	γ_1	1,7
ζ	1	ρ	3

Таблица 3. Начальные значения переменных.

Переменные	Начальные значения	Переменные	Начальные значения
θ_0^k, θ_i^k	Равномерно распределены в $(0, \pi)$	ω_0^k	6
a_0^k	10	a_i^k	2

Литература

1. Казанович Я.Б., Борисюк Р.М. (2005) Осцилляторная нейросетевая модель слежения за движущимся объектом. Нейроинформатика-2005. VII Всероссийская научно-техническая конференция, т. 1. Москва, МИФИ. с. 25-35.
2. Казанович Я.Б., Шматченко В.В. (2004). Осцилляторные нейросетевые модели сегментации изображений и зрительного внимания. Нейроинформатика-2004. VI Всероссийская научно-техническая конференция. Лекции по нейроинформатике, т. 1, под ред. Ю.В. Тюменцева. Москва: МИФИ, с. 72-125.
3. Крюков В.И. (2004). Модель внимания и памяти, основанная на принципе доминанты и компараторной функции гиппокампа. Журнал высшей нервной деятельности, 54(1), 10-29.
4. Andres P. (2003). Frontal cortex as the central executive: time to revise our view. Cortex, 39, 871-895.
5. Baddeley A. (1996). Exploring the central executive. Quarterly Journal of Experimental Psychology, 49A, 5-28.
6. Baddeley A. (2002). Fractionating the central executive. In D. Stuss and R. T. Knight (Eds.), Principles of Frontal Lobe Function (pp. 246–260). New York: Oxford University Press.
7. Baddeley A. (2003). Working memory and language: an overview. Journal of Communication Disorders, 36, 189-208.

8. Barbas H. (2000). Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. *Brain Research Bulletin*, 52, 319–330.
9. Blaser E., Pylyshyn Z.W., Holcombe A.O. (2000). Tracking an object through feature space. *Nature*, 408, 196-199.
10. Borisyuk R., Kazanovich Y. (2003). Oscillatory neural network model of attention-focus formation and control. *BioSystems*, 71, 29-36.
11. Borisyuk R., Kazanovich Y. (2004). Oscillatory model of attention-guided object selection and novelty detection. *Neural Networks*, 17, 899-915.
12. Collett F., Van der Linden M. (2002). Brain imaging of the central executive component of the working memory. *Neuroscience and Behavioral Review*, 26, 105-125.
13. Corchs S., Deco G. (2001). A neurodynamical model for selective visual attention using oscillators. *Neural Networks*, 14, 981-990.
14. Cowan N. (1988). Evolving conceptions of memory storage, selective attention and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163-191.
15. Culham J., Brandt S.A., Cavanagh P., Kanwisher N.G., Dale A.M., Tootell R. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology*, 80, 2657-2670.
16. Damasio A. (1989). The brain binds entities and events by multiregional activation from convergent zones. *Neural Computation*, 1, 123-132.
17. D'Esposito M., Detre J.A., Alsop D.C., Shin R.R., Atlas S., Grossman M. (1991). The neural basis of central executive system of working memory. *Nature*, 378, 279-281.
18. Duncan J. (2001). An adaptive coding model of neural functions in prefrontal cortex. *Nature Reviews Neuroscience*, 2, 820-829.
19. Egeth H, Yantis S. (1997). Visual attention: control, representation, and time course. *Annual Review of Psychology*, 48, 269-297.
20. Eriksen C.W., St. James J.D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Perception and Psychophysics*, 40, 225-240.
21. Fell J., Fernandez G., Klaver P., Elger C.E., Fries P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Research Reviews*, 42, 265-272.
22. Fries P., Reynolds J.H., Rorie A.E., Desimone R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291, 1560-1563.
23. Fries P., Schroeder J-H., Roelfsema P.R., Singer W., Engel A.K. (2002). Oscillatory neural synchronization in primary visual cortex as a correlate of stimulus selection. *Journal of Neuroscience*, 22, 3739-3754.
24. Gross J., Scmitz F., Schnitzler I., Kessler K., Shapiro K., Hommel B., Schnitzler A. (2004). Modulation of long-range neuronal synchrony reflects temporal limitations of visual attention in humans. *Proc. Natl. Acad. Sci. (USA)*, 101, 13050-13055.
25. Grossberg S., Raizada R. (2000). Contrast sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, 40, 1413-1432.
26. Herrmann C.S., Knight R.T. (2000). Mechanisms of human attention: event related potentials and oscillations. *Neuroscience and Biobehavioral Reviews*, 25, 465-476.
27. Holscher C. (2003). Time space and hippocampal functions. *Review of Neuroscience*, 14, 253-284.
28. Itti L., Koch C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40, 1489-1506.
29. Itti L., Koch C. (2001). Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2, 194-203.
30. Itti L., Koch C., Niebur E. (1998). A model of saliency-based visual attention for rapid screen analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 1254-1259.

31. Kanwisher N., Wojciulik E. (2000). Visual attention: Insights from brain imaging. *Nature Reviews Neuroscience*, 1, 91-100.
32. Katayama K., Yano M., Horiguchi T. (2004). Neural network model of selective visual attention using Hodgkin-Huxley equation. *Biological Cybernetics*, 91, 315-325.
33. Kazanovich Y.B., Borisjuk R.M. (1999). Dynamics of neural networks with a central element. *Neural Networks*, 12, 441-454.
34. Kazanovich Y., Borisjuk R. (2002). Object selection by an oscillatory neural network. *BioSystems*, 67, 103-111.
35. Kazanovich Y.B., Borisjuk R.M. (2003). Synchronization in oscillator systems with phase shifts. *Progress in Theoretical Physics*, 110, 1047-1058.
36. Koch C., Ullman S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4, 219-227.
37. Kryukov V.I. (1991). An attention model based on the principle of dominance. In A.V. Holden and V.I. Kryukov (Eds.), *Neurocomputers and Attention I. Neurobiology, Synchronization and Chaos* (pp. 319-352). Manchester: Manchester University Press.
38. Kuramoto Y. (1991). Collective synchronization of pulse coupled oscillators and excitable units. *Physica D*, 50, 15-30.
39. Liang H., Bressler S.L., Ding M., Desimone R., Fries P. (2003). Temporal dynamics of attention-modulated neuronal synchronization in macaque V4. *Neurocomputing*, 52-54, 481-487.
40. Liu G., Austen E.L., Booth K.S., Fisher B.D., Argue R., Rempel M.I., Enns J.T. (2005). Multiple-object tracking is based on scene, not retinal coordinates. *Journal of Experimental Psychology*, 31, 235-247.
41. Loose R., Kaufmann C., Auer D.P., Lange K.W. (2003). Human prefrontal and sensory cortical activity during divided attention tasks. *Human Brain Mapping*, 18, 249-59.
42. Moser M.C., Siltan M. (1998). Computational model of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341-393). London: UCL Press.
43. Motter B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, 70, 909-919.
44. Niebur E., Hsiao S.S., Johnson K.O. (2002). Synchrony: a neuronal mechanism for attentional selection? *Current Opinion in Neurobiology*, 12, 190-194.
45. Niebur E., Kammen D.E., Koch C. (1991). Phase-locking in 1-D and 2-D networks of oscillating neurons. In W. Singer and H. Schuster (Eds.) *Nonlinear Dynamics and Neuronal Networks*, (pp. 173-204). Vieweg Verlag, Berlin.
46. Oksama L., Hyönä J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, 11, 631-671.
47. Olshausen B.A., Anderson C.H., Van Essen D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13, 4700-4719.
48. Posner M.I., Snyder C.R.R., Davidson D.J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160-174.
49. Pylyshyn Z.W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80, 127-158.
50. Pylyshyn Z. W., Storm R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial Vision*, 3, 179-197.
51. Roelfsema P.R., Lamme V., Spekreijse H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395, 376-381.
52. Scholl B.J. (2001). Objects and attention: the state of the art. *Cognition*, 80, 1-46.

Нейросетевая модель слежения за несколькими объектами

53. Scholl B.J., Tremoulet P.D. (2000). Perception causality and animacy. *Trends in Cognitive Science*, 4, 299-309.
54. Sears C.R., Pylyshyn Z.W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology*, 54, 1-14.
55. Seifert A.E., Somers D.C., Dale A.M., Tootel R. (2003). Functional MRI studies of human visual motion perception: texture, luminance, attention and aftereffects. *Cerebral Cortex*, 13, 340-349.
56. Shallice T. (2002). Fractionation of the supervisory system. In D. T. Stuss and R. T. Knight (Eds.), *Principles of Frontal Lobe Function* (pp. 261–277). New York: Oxford University Press.
57. Singer W. (1999). Neuronal synchrony: A versatile code for the definition of relations. *Neuron*, 24, 49-65.
58. Sokolov A., Lutzenberger W., Pavlova M., Pressl H., Braun C., Birbauner N. (1999). Gamma-band MEG activity to coherent motion depends on task-driven attention. *Neuroreport*, 10, 1997-2000.
59. Somers, D.C., Dale A.M., Seifert A.E., Tootel R. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proc. Natl. Acad. Sci. (USA)*, 96, 1663-1668.
60. Steinmetz P.N., Roy A., Fitzgerald P., Hsiao S.S., Johnson K.O., Niebur E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, 404, 187-190.
61. Tallon-Baudry C. (2004). Attention and awareness in synchrony. *Trends in Cognitive Sciences*, 8, 523-525.
62. Treue S. (2003). Visual attention: the where, what, how and why of salience. *Current Opinion in Neurobiology*, 13, 428-432.
63. Tsotsos J.K., Culhane S.M., Wai W.Y.K., Lai Y., Davis N., Nufl F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, 78, 507-545.
64. Vinogradova O.S. (2001). Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*, 11, 578-598.
65. Visvanathan L., Mingolla E. (2002). Dynamics of attention in depth: evidence from multi-element tracking. *Perception*, 31, 1415-1437.
66. Wang D.L. (1999). Object selection based on oscillatory correlation. *Neural Networks*, 12, 579-592.
67. Yantis S. (1992). Multielement visual tracking: attention and perceptual organization. *Cognitive Psychology*, 24, 295-340.

Статья поступила 10 ноября 2005 г.