

УДК: 001(06)+004.032.26(06)

Теория временной корреляции и модели сегментации зрительной информации в мозге (обзор)

Казанович Я.Б.*

*Институт математических проблем биологии, Российская академия наук, Пущино,
Московская область, 142290, Россия*

Аннотация. Обзор посвящен описанию различных подходов к построению моделей сегментации изображений на основе использования динамических процессов, колебаний и синхронизации нейронной активности. Приводятся нейробиологические основы моделей, описываются принципы функционирования моделей и результаты их применения. Обсуждаются сравнительные характеристики моделей и перспективы их улучшения.

Ключевые слова: сегментация изображений, осцилляторные нейронные сети, синхронизация, интеграция признаков, память, внимание.

СОДЕРЖАНИЕ

1. Введение. Теория временной корреляции	45
2. Сегментация контурных изображений	52
2.1. Модель Ли	52
2.2. Модель Чое-Миккулаинена.....	55
2.3. Модель Ю-Жанга	58
2.4. Обсуждение	59
3. Сегментация пиксельных изображений	59
3.1. Параллельные процедуры.	60
3.2. Последовательные процедуры.	63
3.3. Обсуждение	69
4. Сегментация и кластеризация	70
4.1. Модель Роума-Фригуи	70
4.2. Модель Жао	74
4.3. Модель Буманна-Ланге-Рамахера	77
4.4. Обсуждение	79
5. Сегментация и когнитивные функции мозга	80
5.1. Сегментация и память	80
5.2. Сегментация и внимание	83
5.3. Обсуждение	91
6. Заключение	92
Список литературы	93

1. ВВЕДЕНИЕ. ТЕОРИЯ ВРЕМЕННОЙ КОРРЕЛЯЦИИ

Разбиение изображений на непересекающиеся фрагменты, соответствующие зрительным объектам и фону, является популярной задачей искусственного интеллекта. На эту тему опубликовано множество работ. Обзоры по данной области исследований можно найти, например, в работах [25, 45, 54, 56, 65, 66, 82, 86] (см. также веб-страницы [78, 89]). Сегментация считается важным предварительным шагом, существенно облегчающим последующее решение задач анализа зрительных сцен, таких как распознавание образов, запоминание изображений, обнаружение заданных объектов и т.д. Хотя люди используют достаточно единообразные интуитивные стратегии сегментации, дать формальное, универсальное и объективное определение качества сегментации вряд ли возможно. Многочисленные эксперименты показывают, что сегментация зависит от контекста, предшествующего опыта и внутренних установок, которые лежат за пределами самого анализируемого изображения. Все это препятствует формулировке сегментации как чисто оптимизационной задачи.

Компьютерная реализация сегментации опирается в значительной мере на интуицию и здравый смысл. Обычно применяющиеся методы основаны на предположении, что пиксели в сегментированной области должны иметь сходные оптические характеристики по интенсивности, цвету, текстуре, в то время как разные области хотя бы по одной из этих характеристик различаются. Кроме того, сегментация должна быть устойчива по отношению к шуму и естественной вариации объектов и фона. Предпочтение отдается тем методам, которые легче адаптируются к широкому классу изображений путем соответствующей модификации параметров алгоритма.

В то время как люди в большинстве случаев успешно справляются с задачей сегментации, часто даже на бессознательном уровне, компьютерная сегментация оказалась трудной задачей. В связи с этим у исследователей возникло желание лучше разобраться, как эту задачу решает мозг. Важной особенностью работы мозга является параллельная обработка информации. Это дает надежду на то, что использование биологических методов сегментации окажется не только качественным, но и быстрым.

Встречный интерес к проблеме сегментации возник и у биологов в связи с так называемой проблемой интеграции признаков объектов в цельный образ. В англоязычной литературе эта проблема называется проблемой байндинга (the binding problem – проблема связывания; в дальнейшем термин байндинг будет использоваться как синоним интеграции признаков). Дело в том, что признаки, используемые при кодировании информации о зрительных стимулах в первичных зонах зрительной коры, имеют различную природу. Это могут быть геометрические, спектральные или динамические характеристики изображения. Известно, что первичная обработка различных по своей природе признаков идет в специализированных нейронных структурах коры мозга. Только в ассоциативных зонах коры формируется представление о целостных объектах. В связи с этим возникает вопрос: каковы нейронные механизмы, позволяющие сохранить информацию о принадлежности признаков к отдельным объектам, и как осуществляется объединение нужных признаков в цельный образ объекта? Как писал один из пионеров в исследованиях байндинга фон дер Мальсбург: "Разум – это место встречи и источник всех символических систем. Нет сомнения, что он должен обладать механизмом для комбинирования своих семантических атомов в молекулы и агрегаты, и этот механизм богаче всего, что было до сих пор придумано в теории или реализовано в электронике. Понять, в чем заключается этот механизм – значит решить проблему байндинга".

Нетривиальность проблемы байндинга становится очевидной при одновременном предъявлении на изображении нескольких объектов. В этом случае весь набор признаков, извлеченных из изображения, должен быть сегментирован с указанием того,

какой признак какому объекту принадлежит, и "метка" принадлежности должна сохраняться на всем пути продвижения зрительной информации по коре. Имеются психологические эксперименты, показывающие, что при недостатке ресурсов (например, нехватке времени) испытуемые могут выполнять группировку признаков по объектам с ошибками. Например, после предъявления изображения, содержащего зеленый треугольник и красный квадрат, испытуемому может показаться, что ему были показаны зеленый квадрат и красный треугольник.

Из сказанного видно, что сегментацию можно рассматривать как частный случай байндинга. В самом деле, если речь идет об изображениях, заданных попиксельно, то достаточно включить в число признаков объектов признаки отдельных пикселей, чтобы метка принадлежности пикселей различным объектам реализовывала сегментацию изображения. Если же изображение задано в виде контуров объектов, тогда для сегментации объектов достаточно, чтобы в число признаков были включены признаки сегментов, составляющих контуры.

Гипотеза о возможном механизме интеграции признаков в мозге была предложена в работе [47]. Основная идея состоит в том, что признаки объекта кодируются синфазной активностью нейронов в различных областях коры мозга. Синфазность служит в качестве метки, помечающей информацию, относящуюся к одному объекту. Таким образом, информация, относящаяся к одному зрительному объекту, должна кодироваться синфазной активностью нейронов по всему пути обработки этой информации от первичных зон коры до ассоциативных и моторных зон. Информация, относящаяся к разным объектам, одновременно присутствующим в зрительной сцене, должна кодироваться асинхронной (некоррелированной) активностью между ансамблями нейронов, представляющими данные объекты. Эта гипотеза легла в основу теории интеграции признаков, получившей название теории временной корреляции (ТВК) [22, 67, 68]. Ее частным случаем, когда нейронная активность носит четко выраженный колебательный характер, является теория осцилляторной корреляции (oscillatory correlation theory) [74].

Первые экспериментальные данные, связанные с ТВК, были получены в работах [16, 23]. В экспериментах на первичных зонах зрительной коры животных была обнаружена как стимулоспецифическая колебательная активность нейронов (рис. 1), так и синфазность этой активности у нейронов, кодирующих один и тот же стимул. В частности, проводились эксперименты на кошках, которым предъявлялась светлая полоска, движущаяся по темному экрану. Активность на уровне одиночных нейронов (multiunit activity) и нейронных ансамблей (local field potential) регистрировалась двумя электродами, установленными в первичных (стриарных) зонах зрительной коры.

Были получены следующие результаты.

1. Нейроны отвечают периодической активностью с частотой в гамма-диапазоне (40-70 герц) при прохождении полоски через их рецептивное поле.
2. При одновременном пересечении полоской двух рецептивных полей активность соответствующих нейронов синфазна, даже если эти нейроны находятся друг от друга на расстоянии нескольких миллиметров.
3. Если полоску разорвать на две изолированных части, то синхронность ответа падает тем больше, чем больше величина разрыва.
4. Если на экране движутся две полоски в разных направлениях (одна вверх, другая вниз), то синхронизованного ответа не возникает даже в тот момент, когда эти полоски оказываются на одной прямой.

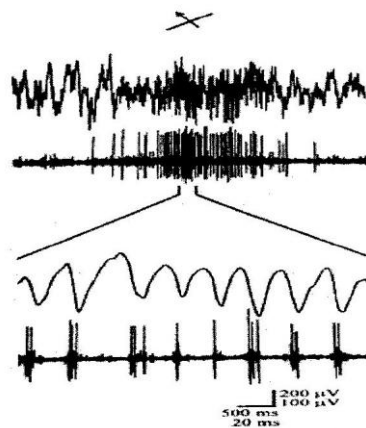


Рис. 1. Локальный полевой потенциал и множественная нейронная активность, записанные в поле 17 зрительной коры кошки в ответ на предъявление светлой полоски с оптимальной ориентацией, перемещающейся через рецептивное поле записываемых нейронов. Два нижних графика показывают те же данные для уменьшенного временного интервала. Отметим высокую степень синхронизации между импульсами нейронной активности и минимумами локального полевого потенциала [21].

Таким образом, в случае 2 имеет место синхронный ответ на единичный объект, а в случаях 3-4 несинхронный ответ на два разных объекта.

В дальнейшем были проделаны многочисленные эксперименты на разных животных и с многоэлектродной записью активности различных участков мозга, подтвердившие важную роль осцилляторной и синхронной активности в реализации когнитивных функций. Пример, иллюстрирующий эти экспериментальные данные, приведен на рис. 2. Запись велась двумя электродами в полях 17 (стриарная кора – первичная зона зрительной коры) и PMLS (posteromedial lateral suprasylvian area – ассоциативная зрительная область, специализирующаяся на анализе движения) Множественные пики на рис. 2Г указывают на наличие колебательной активности в регистрируемых сигналах. Хорошо выраженный пик в районе нуля в левой части рис. 2Д показывают высокий уровень синхронизации (синфазности) активности, регистрируемой в нейронах при предъявлении стимула в виде сплошной полоски. Величина этого пика уменьшается при предъявлении разорванной полоски (средняя часть рис. 2Д). Корреляция полностью исчезает при предъявлении двух полосок, движущихся в противоположных направлениях (правая часть рис. 2Д).

Реализация байндинга на уровне первичных зон зрительной коры возможна лишь для простых стимулов типа движущихся полосок. Восприятие сложных зрительных стимулов требует взаимодействия между различными уровнями обработки внешних стимулов, вниманием и памятью [51]. Внимание используется для оптимальной организации этого взаимодействия. В частности, показано, что внимание играет важную роль в байндинге, давая возможность связывать неоднотипные признаки (включая признаки разной модальности).

Экспериментальные данные, проясняющие роль синхронизации нейронной активности в реализации памяти и внимании, стали появляться, лишь в последнее десятилетие. [30, 61, 64] (см. также обзор [26]). Считается, что синхронизация нейронной активности создает подходящие условия для долговременной синаптической потенциации, являющейся основным кандидатом на механизм формирования кратковременной памяти [80]. В работе [48] это предположение подтверждено путем анализа ЭЭГ при ассоциативном обучении у людей. Это обучение приводило к повышению когерентности гамма-колебаний в двух областях коры, получающих два класса стимулов, между которыми должна быть установлена ассоциация. В работе [64] было показано, что хорошее запоминание слов людьми

связано с повышением мощности тета- и гамма-активности в определенных участках коры. В работе [30] было обнаружено повышение синхронизации на частоте гамма-ритма в гиппокампе обезьян при решении задачи узнавания запомненных зрительных образов.

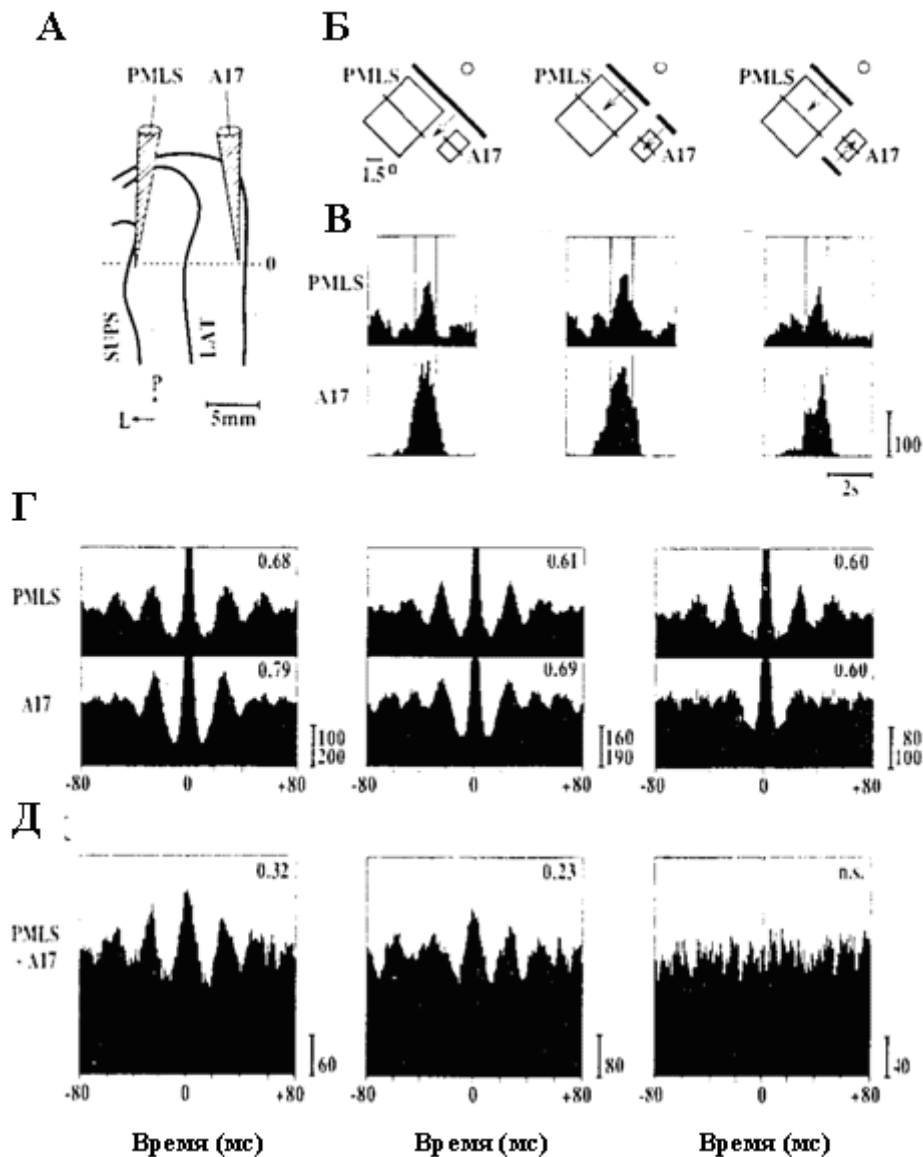


Рис. 2. Синхронизация нейронной активности между полями A17 (area 17) и PMLS (posteromedial lateral suprasylvian area) кошки при воздействии различными зрительными стимулами. **А.** Расположение электродов. LAT – lateral external sulcus (латеральная внешняя борозда), SUPS – suprasylvian sulcus (супрасильвиева борозда), P – постериорное, L – латеральное направления. **Б.** Типы стимулов и расположение рецептивных полей в A17 и PMLS. Маленьким кружком обозначен центр поля зрения. **В.** Постстимульные гистограммы для трех рассматриваемых стимулов. **Г.** Автокорреляционные функции для трех видов стимулов. **Д.** Кросскорреляционные функции для трех видов стимулов [21].

Синхронизация нейронной активности при запоминании информации может быть соотнесена с процедурой сравнения стимула, находящегося в фокусе внимания, с содержимым рабочей памяти [62]. Тот факт, что функционирование внимания связано с осцилляторной активностью и ее синхронизацией был установлен в работах [15, 19, 20, 69]. Однако до недавнего времени не было известно, какие структуры мозга участвуют в реализации синхронной активности в фокусе внимания, Постепенно накапливаются экспериментальные и теоретические основания полагать, что система внимания

является иерархической, то есть что в ней имеется некоторая подсистема, называемая центральным элементом, которая организует сегментацию определенного объекта и включение его в фокус внимания [1, 2, 14]. В работах [38, 93] была выдвинута гипотеза о том, что взаимодействие центрального элемента системы внимания с нейронными ансамблями, представляющими различные объекты, происходит по принципу фазовой автоподстройки частоты. В фокус внимания включается тот объект, признаки которого кодируются в коре осцилляторной активностью, синхронизованной с активностью центрального элемента.

Недавно эта гипотеза получила экспериментальное подтверждение. В работе [24] было показано, что при внимании определенная область в префронтальной коре – фронтальное глазное поле FEF (frontal eye field) - является источником повышенной синхронизации нейронной активности на частоте гамма-ритма в поле V4.

Эксперимент был организован следующим образом (рис. 3). В начале эксперимента обезьяна фиксировала взгляд в центре экрана и нажимала на педаль, сигнализируя о готовности. После этого ей предъявлялось изображение с тремя стимулами (круги разного цвета - красный, зеленый, синий - с черными полосками). Внимание к целевому стимулу направлялось с помощью подсказки - центр фиксации окрашивался в цвет выбираемого стимула (в случае, представленном на рисунке, это красный цвет). Через некоторое время окраска целевого стимула светлела и обезьяна фиксировала этот факт отпуская педаль.

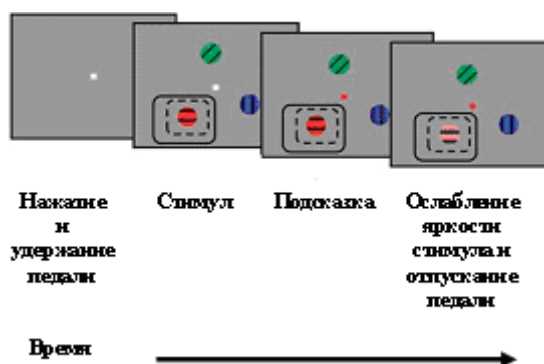


Рис. 3. Схема эксперимента. Прямоугольники, нарисованные пунктиром и сплошной линией вокруг красного стимула иллюстрируют гипотетические рецептивные поля нейронов в полях V4 и FEF. [24].

Активность записывалась двумя электродами. Каждый электрод мог регистрировать либо импульсную множественную нейронную активность (МНА), либо локальный полевой потенциал (ЛПП). Рассматривались следующие варианты установки электродов: 1) оба электрода в V4, 2) оба электрода в FEF., 3) один электрод в V4, а другой в FEF.

Вычислялся некоторый коэффициент, который характеризует степень когерентности (синхронизованности) сигналов, записанных по указанным двум каналам. Результаты экспериментов суммированы на рис. 4 в виде графиков, показывающих зависимость когерентности от частоты. Общий вывод – внимание приводит к повышению синхронизации активности между полями V4 и FEF в диапазоне гамма ритма 40 - 80 гц (наибольшее увеличение синхронизации наблюдалось в районе 50 гц). В диапазоне тета- и бета-ритмов (5 - 25 гц) изменение степени синхронизованности в связи с фиксацией внимания на стимуле не наблюдалось.

Психологические исследования показывают, что внимание улучшает качество сегментации изображений. Это порождает надежду на то, что тот же положительный эффект может быть получен при включении внимания в модели сегментации. Общей идеей здесь является соединение принципов синхронизации и модуляции нейронной

активности. Идея центрального элемента системы внимания также является плодотворной для моделирования, так как позволяет заменить большое число глобальных связей между элементами модели на связи с центральным элементом.

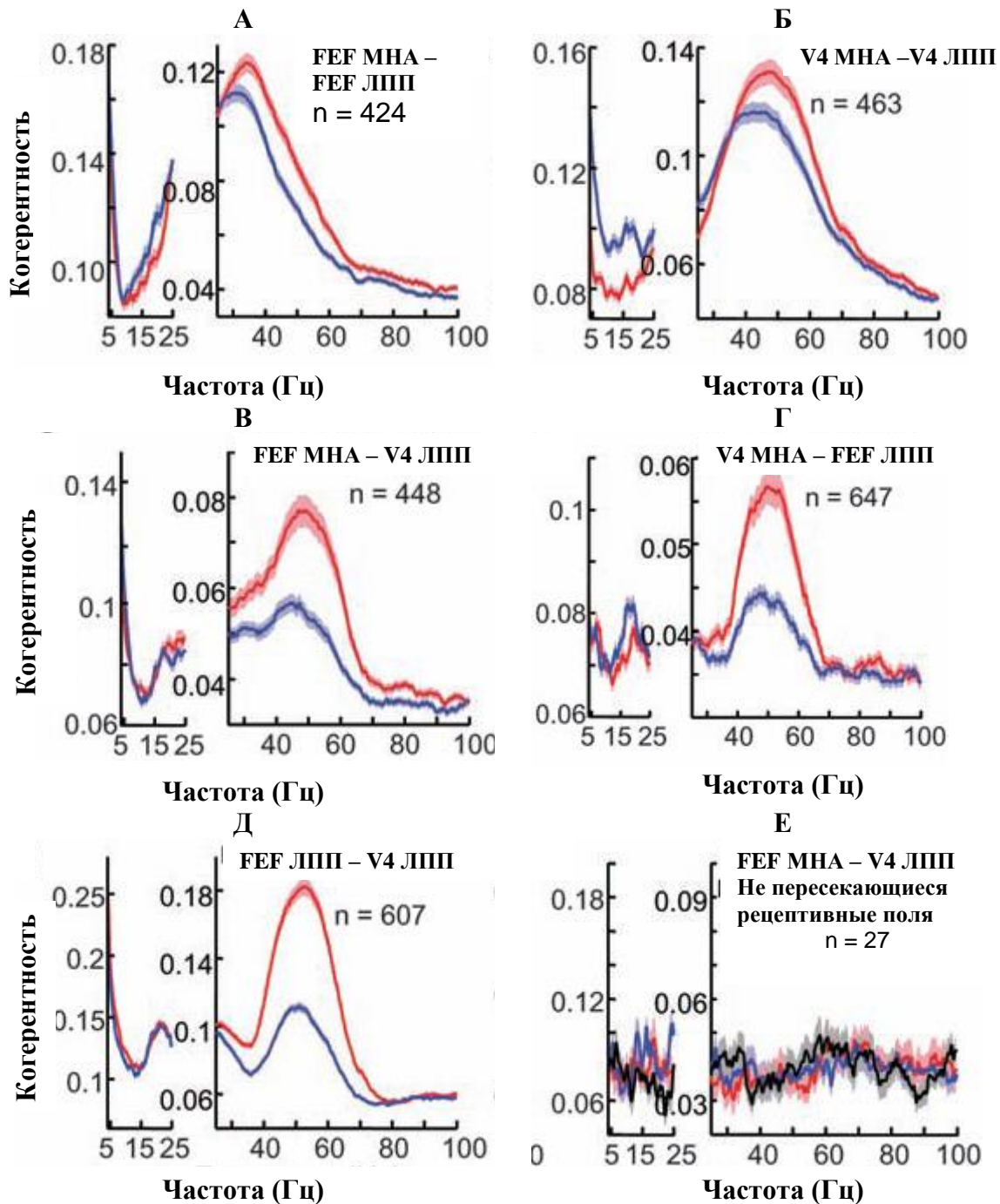


Рис. 4. Повышение синхронизации при внимании между полями V4 и FEF MHA – множественная нейронная активность, ЛПП – локальный полевой потенциал. А - Д. Графики зависимости когерентности от частоты в случае, когда оба электрода регистрируют активность нейронов с пересекающимися рецептивными полями, реагирующих на стимул в фокусе внимания (красные кривые) и на стимул вне фокуса внимания (синие кривые). Существенное улучшение синхронизации имеет место на рис. В - Д в окрестности частоты 50 гц. Е. Графики зависимости когерентности от частоты в случае, когда оба электрода регистрируют активность нейронов с непересекающимися рецептивными полями, Внимание сфокусировано на рецептивном поле V4 (красная кривая), внимание сфокусировано вне рецептивных полей V4 и FEF (синяя кривая), внимание сфокусировано в центре поля зрения при отсутствии стимулов (черная кривая). Синхронизации активности между полями V4 и FEF в данных случаях не наблюдается [24].

Несмотря на то, что ТВК является наиболее продвинутой попыткой объяснения интеграции признаков, следует отметить, что данная теория не может считаться общепризнанной, а эксперименты, подтверждающие ТВК, поставлены при весьма специальных условиях: используемые в экспериментах стимулы, лишь очень отдаленно напоминают реальные зрительные объекты. Однако, поскольку принципы ТВК кажутся перспективными для обработки изображений, теоретическое и модельное изучение роли синхронизации нейронной активности в обработке информации пошло намного дальше имеющихся экспериментальных данных. В частности, целью исследований является проверка, каким образом ТВК может быть применена для анализа сложных изображений с нечетко определенными границами между объектами и высоким уровнем шумов. В случае успеха эти результаты могут быть полезны и для нейробиологов, поскольку помогут им в поиске новых экспериментальных фактов для обоснования ТВК. С другой стороны, ТВК может оказаться весьма эффективным инструментом при решении практических задач анализа изображений, поскольку принципы ТВК достаточно просто могут быть воплощены в технических системах с параллельными вычислениями.

Идеология ТВК хорошо сочетается с методами моделирования, опирающимися на осцилляторные нейронные сети [87, 88]. Согласно ТВК сегментация должна приводить к синфазной активности нейронов, кодирующих зрительный объект, в то время как активность нейронных ансамблей, кодирующих разные объекты, должна быть некогерентной. Общая идея, присутствующая в большинстве нейросетевых моделей сегментации, состоит в реализации синфазности с помощью локальных латеральных связей (такая организация связей согласуется с архитектурой связей в первичных зонах зрительной коры). Другая идея, являющаяся обобщением имеющихся нейрофизиологических данных, заключается в предположении, что входное изображение кодируется в нейронной сети таким образом, что соседние элементы сети (нейроны или осциллирующие нейронные ансамбли) имеют тенденцию работать синфазно лишь в том случае, если участки изображения, находящиеся в их рецептивных полях, обладают сходными оптическими характеристиками. Опираясь при построении модели на эти предположения, можно ожидать, что динамика нейронной сети приведет к формированию в сети синфазных кластеров нейронов (при отсутствии синфазности или даже синхронизации между кластерами), таких что соответствующий каждому кластеру сегмент изображения будет иметь сходные (или, по крайней мере, плавно меняющиеся) оптические характеристики.

В последние годы было предложено большое количество моделей сегментации, отличающихся степенью учета биологических данных, классами анализируемых изображений, применяемыми принципами функционирования и результатами применения. Обзор ранних результатов по этой тематике можно найти в работе [92]. Некоторые авторы обращают внимание на контурные изображения, поскольку они более похожи на стимулы, используемые в нейрофизиологических экспериментах. Другие предпочитают полутоновые и цветные изображения в силу их большей практической значимости. В последнем случае рецептивные поля нейронов или нейронных осцилляторов обычно представлены пикселями на изображении, каждый из которых характеризуется интенсивностью серого (для полутоновых изображений) или интенсивностями трех-четырёх составляющих спектра (для цветных изображений).

Принятый в данной работе подход к категоризации методов сегментации является компромиссом, старающимся учесть как тип обрабатываемых изображений, так и применяемые вычислительные схемы. Сначала будут рассмотрены модели для сегментации контурных изображений. Затем для изображений, заданных попиксельно, описываются два класса моделей сегментации – с последовательным и параллельным выбором объектов. Результатом работы параллельной процедуры является

одновременное существование нескольких синфазных внутри себя нейронных ансамблей, каждый из которых кодирует свой зрительный объект. При использовании последовательной процедуры синфазные ансамбли, кодирующие различные объекты, возникают один за другим, так что в каждый момент времени только один из представленных одновременно в зрительной сцене объектов кодируется когерентной активностью определенного нейронного ансамбля, в то время как активность в остальной части сети отсутствует или подавлена до низкого уровня. Далее будет рассмотрено, как сегментация может сочетаться с методами кластеризации. В следующем разделе описываются модели сегментации, привлекающие дополнительные когнитивные функции, такие как память и внимание. Заключительный раздел посвящен общему обсуждению существующих моделей сегментации и перспективам их совершенствования.

Терминологическое замечание. В работе часто будут использоваться термины **синхронизация**, **синфазность**, **когерентность**, **противофазность**. Под синхронизацией двух сигналов будет пониматься ситуация, когда на кросскорреляционной функции этих сигналов имеется хотя бы один большой пик (в случае периодической активности таких пиков будет несколько). Когерентность будет использоваться как синоним синхронности. Если основной пик кросскорреляционной функции находится в нуле, будем говорить о синфазной активности. Если речь идет об импульсной активности, синфазность означает, что импульсы в обоих сигналах генерируются одновременно. Противофазному режиму соответствует расположение основного пика кросскорреляционной функции в $\pm\pi$. В этом случае моменты максимума одного сигнала соответствуют минимумам другого сигнала. Под нейронным **ансамблем** будем понимать связанное множество синфазно работающих нейронов, представляющих один из объектов на изображении. Аналогичное соглашение имеет место для ансамблей нейронных осцилляторов.

2. СЕГМЕНТАЦИЯ КОНТУРНЫХ ИЗОБРАЖЕНИЙ

Сегментация контурных изображений является специальным случаем задачи сегментации, когда входное изображение задано в виде набора коротких прямолинейных отрезков, имеющих различную ориентацию и располагающихся в узлах двумерной решетки. Психологические и нейрофизиологические наблюдения показывают, что интеграция признаков в таких изображениях происходит в первичных зонах зрительной коры (стриарная кора) в основном благодаря латеральному взаимодействию нейронов, имеющих сходные ориентационные предпочтения (т.е. дающих максимальную реакцию на стимулы примерно одной и той же ориентации). Это, впрочем, не исключает определенной роли обратных связей от вышележащих корковых областей (в частности, экстрастриарных). Целью моделирования контурной сегментации является разработка такой биологически правдоподобной нейронной сети, которая бы позволяла осуществить группировку отрезков в контуры при подавлении активности нейронов, вызванной дистракторами (ложными объектами) и шумом. Дополнительно к этому модель должна обладать способностью заполнять пробелы в неполных контурах с тем, чтобы результирующий контур был связанным.

2.1. Модель Ли

Модель, описывающая функционирование чувствительных к ориентации нейронов в области V1, была предложена и изучена в серии статей З. Ли [41 - 44]. Параметры модели (кроме входных сигналов) являются жестко заданными в том смысле, что они не зависят от предъявленного изображения и не меняются в процессе его обработки. Параметры связи между парой нейронов зависят лишь от расстояния между ними (сила связи убывает с расстоянием) и от разницы в их предпочтительной ориентации.

Основное предположение модели состоит в том, что два нейрона связаны наибольшей возбуждающей связью, если их предпочтительные ориентации коллинеарны отрезку, соединяющему центры рецептивных полей этих нейронов. Наоборот, в случае ортогональности этих направлений связь между нейронами максимально ингибирующая. Остальные силы связи имеют промежуточные значения. Таким образом, нейроны, активируемые гладко меняющимся контуром, взаимно увеличивают свою активность, в то время как активность нейронов, активируемых шумом, подавляется. Экспериментальное подтверждение такой схемы взаимодействия содержится в работе [32].

Базовым функциональным элементом модели является осциллятор, формируемый парой из возбуждающего и тормозного нейронов. Динамика осциллятора описывается уравнениями

$$\begin{aligned}\frac{dx}{dt} &= -x - g(y) + J_0 f(x) + I_x, \\ \frac{dy}{dt} &= -y + f(x) + I_y.\end{aligned}$$

Здесь переменные x и y задают мембранные потенциалы возбуждающего и тормозного нейронов, соответственно, f и g – кусочно-линейные сигмоидные функции с порогом и насыщением, J_0 – сила синаптической связи между возбуждающими клетками, I_x, I_y – входные сигналы. В зависимости от входных сигналов и параметров осциллятор может демонстрировать различные виды динамики, включающие стационарные состояния, регулярные и хаотические колебания.

Возбуждающие нейроны в модели имеют рецептивные поля, чувствительные к ориентации стимула. Другими словами, внешний сигнал, приходящий на нейрон, максимален, если ориентация отрезка на изображении совпадает с предпочтительной ориентацией клетки. Этот сигнал убывает по мере того, как растет рассогласование между этими направлениями. Кроме того, входной сигнал увеличивается в случае отрезка, обладающего большей заметностью (то есть отличающегося по направлению от окружающих его отрезков).

В каждом узле двумерной решетки находится не один осциллятор, а гиперколонка, состоящая из K осцилляторов с предпочтительными ориентациями, равными $\theta = k\pi / K$ ($k = 0, \dots, K - 1$) (при вычислениях использовалось 12 направлений, $K = 12$).

В модели реализована следующая архитектура связей. Внутри гиперколонки имеются связи от каждого тормозного нейрона на все возбуждающие нейроны. Сила связи убывает с ростом $|\Delta\theta|$ (максимальная сила соответствует связи, идущей на возбуждающий нейрон того же осциллятора). Связи между гиперколонками идут от возбуждающего нейрона одной гиперколонки на все возбуждающие и тормозные нейроны соседних гиперколонок. Сила связи убывает по мере возрастания расстояния между узлами решетки. Наиболее важное свойство этих связей состоит в том, что сила связи между возбуждающими нейронами имеет большую величину, если ориентационная предпочтительность нейронов совпадает с ориентацией линии, соединяющей узлы, где расположены эти нейроны. С другой стороны, связь от возбуждающего нейрона к тормозному имеет наибольшую силу, если ориентационная предпочтительность соответствующих осцилляторов ортогональна линии, соединяющей места их расположения.

Дополнительно к описанным входным сигналам на нейроны подаются еще так называемые нормализующие токи, которые предназначены для получения требуемой динамики сети и для управления заметностью контуров. Предполагается, что эти сигналы поступают из более высоких отделов зрительной коры. В частности, подача

сильного сигнала на тормозные нейроны осцилляторов, соответствующих некоторому контуру, может привести к тотальному подавлению активности этих осцилляторов. Результатом будет исключение данного контура из процедуры сегментации.

Основные результаты применения модели (аналитические и компьютерные) были получены для простых геометрических форм, таких как прямые линии и круги в присутствии шума (рис. 5А), но модель применялась и к небольшому числу фотографических изображений, предварительно преобразованных в векторную форму (рис. 5Б). При подходящем выборе параметров модель демонстрировала затухающие или устойчивые синхронные колебания для контуров, представленных отрезками с плавно меняющимися направлениями. Степень выраженности контура в выходных сигналах, идущих от возбуждающих нейронов, и степень синхронизации увеличиваются в случае длинных, гладких и замкнутых контуров, что совпадает с данными биологических экспериментов. Например, набору отрезков, ориентированных вдоль контура окружности, соответствует более высокий выходной сигнал по сравнению с ситуацией, когда отрезки расположены на контуре окружности, но ортогональны к нему.

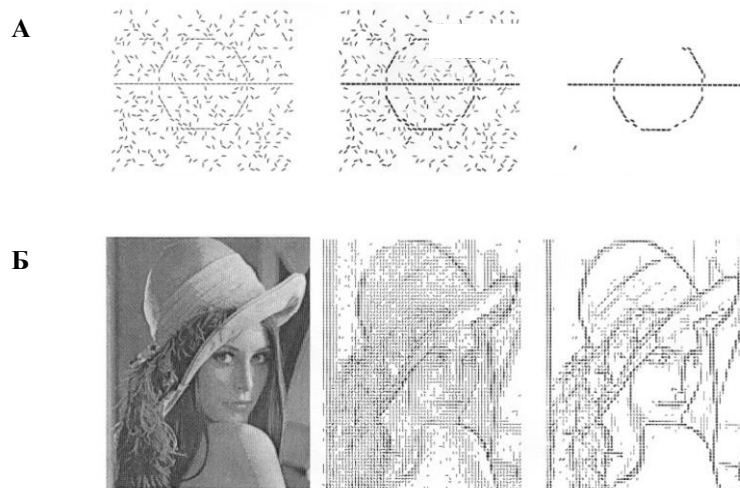


Рис. 5. Примеры обработки контурных изображений моделью Ли. Величина входных и выходных сигналов закодирована толщиной отрезков. В качестве выходного сигнала используются значения $f(x)$, усредненные в течение достаточно большого времени после стимуляции. **А.** Обработка искусственного изображения. Слева – изображение на входе модели. В центре – активность нейронов на выходе модели. Справа – активность нейронов на выходе модели после обрезания активности на уровне 0,63 от максимальной. **Б.** Обработка реального изображения. Слева – исходное изображение. В центре – изображение на входе модели. Справа – активность на выходе модели. [41].

Синхронизация в пределах одного контура имеет место, даже если входные сигналы от отрезков на изображении поступают в модель не одновременно, а с небольшим запаздыванием один после другого. С другой стороны, синхронизация отсутствует, если даже разные контуры пересекаются и все сигналы от изображения приходят одновременно. Последнее вытекает в основном из взаимно подавляющего взаимодействия между осцилляторами в одной гиперколонке.

В работах [42 - 44] основное внимание уделено текстурам, построенным из отрезков с разным наклоном. Показано, что имеет место увеличение выходных сигналов на границе между двумя текстурами с разными наклонами отрезков (так называемый пор-оут границы).

2.2. Модель Чое-Миккулаинена

Хотя модель Ли хорошо соответствует ряду экспериментальных данных, она никак не отвечает на вопрос, каким образом в мозгу возникает требуемая для интеграции контуров архитектура связей. Кроме того, модель не объясняет тот факт, что качество выполнения интеграции падает при перемещении объекта от центра к периферии зрительного поля. В статье [13] предлагается усовершенствованная модель, позволяющая увеличить охват экспериментальных данных на основе зависящей от входных сигналов самоорганизации латеральных связей. Более ранняя версия модели, названная RF-SLISSOM (Receptive Field Spiking Laterally Interconnected Synergetically Self-Organizing Map), была описана в работе [12]. Модель оперирует непосредственно с интенсивностями пикселей без явного использования нейронов с ориентационной селективностью. Никакой заранее заданной упорядоченности при определении величин связей не предполагается. Модель автоматически формирует сильные латеральные связи между нейронами, получающими сигналы от линий со сходной ориентацией.

Базовым элементом модели является интегративно-пороговый импульсный нейрон с утечкой Эхорна [17], работающий в дискретном времени. На каждом синапсе нейрона происходит суммирование приходящих импульсов при экспоненциальном затухании постсинаптического потенциала

$$s(t) = x(t) + s(t-1)e^{-\lambda},$$

где $s(t)$ и $s(t-1)$ - постсинаптические потенциалы в моменты времени t и $t-1$; $x(t)$ - амплитуда входного импульса; λ - параметр затухания. Для разных связей величина λ может быть различной. Подбирая подходящие значения для λ , можно превратить нейрон в детектор совпадения (λ велико) или во временной интегратор (λ мало).

Мембранный потенциал нейрона m формируется как результат действия кусочно-линейной сигмообразной функции на взвешенную сумму постсинаптических потенциалов. Нейрон генерирует импульс в момент времени, когда его мембранный потенциал превышает порог. Величина порога складывается из трех компонент. Первая из них является константой (базовый порог). Вторая задает период абсолютной рефрактерности. Третья реализует относительную рефрактерность и представляет собой функцию, экспоненциально убывающую со временем, прошедшим после последней генерации импульса. При постоянном входном сигнале нейрон работает как осциллятор, частота импульсов которого тем выше, чем больше входной сигнал.

Архитектура связей в модели представлена на рис. 6. Нейронная сеть состоит из двух плоских слоев, называемых картами, MAP1 (136 × 136 нейронов) и MAP2 (54 × 54 нейронов)*. Нейроны в обоих слоях с идентичными координатами образуют кортикальную колонку. Нейроны из одной колонки имеют совпадающие рецептивные поля. Каждый нейрон способен формировать как возбуждающие, так и тормозные сигналы. Входными сигналами на нейрон являются: 1) афферентные возбуждающие сигналы от элементов рецептивного поля (пикселей сетчатки), 2) возбуждающие сигналы от соседних нейронов другой карты, 3) латеральные возбуждающие и тормозные сигналы от соседних нейронов своей карты. Внешние сигналы от сетчатки с размерами 46 × 46 или 72 × 72, поступающие на нейроны обеих карт, варьируют в интервале (0, 1). Рецептивное поле нейрона имеет размер 6 × 6.

Карта MAP1 предназначена для самоорганизации сети и превращения ее в упорядоченную структуру. Нейроны в MAP1 первоначально имеют латеральные возбуждающие связи радиуса 7, порождающие дальний порядок, но этот радиус постепенно сокращается до 3, уточняя локальный порядок. Тормозные связи в MAP1 имеют фиксированный радиус 10.

* Мар по-английски – карта.

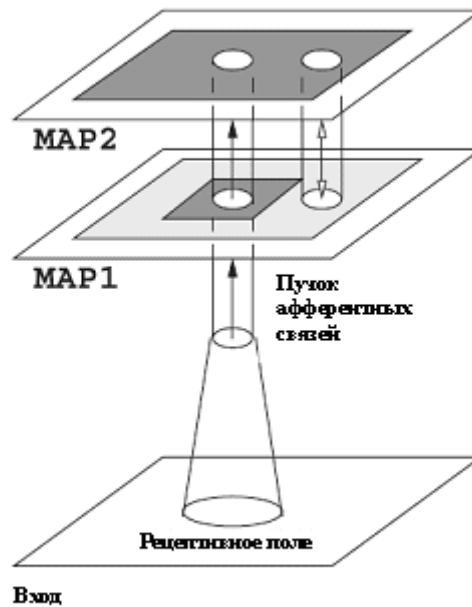


Рис. 6. Архитектура связей модели Чое-Миккулаинена. Односторонние стрелки – восходящие афферентные связи между слоями. Двусторонняя стрелка – связи внутри колонки. Латеральные связи внутри карт: серый цвет – возбуждающие и тормозные связи, белый цвет – тормозные связи [13].

Карта MAP2 предназначена для реализации сегментации и байндинга. В MAP2 возбуждающие латеральные связи имеют радиус 40, а тормозные связи имеют глобальный характер (все-на-всех). Благодаря тормозным связям тенденция к сегментации элементов сети заложена в самой конструкции модели, а для того, чтобы имело место связывание на основе синфазной активности, необходимо действие возбуждающих связей, соответствующая активность в которых порождается входным изображением.

За исключением латеральных тормозных связей в MAP2 веса всех остальных связей модифицируются в соответствии с правилом Хебба с нормализацией

$$w_{ij}(t) = \frac{w_{ij}(t-1) + \alpha V_i(t)V_j(t)}{\sum_k [w_{kj}(t-1) + \alpha V_k(t)V_j(t)]},$$

где w_{ij} - величина связи от нейрона j к нейрону i ; α - параметр, задающий скорость обучения; V_i и V_j - усредненные (в пределах небольшого временного окна) частоты импульсации нейронов. Нормализация препятствует неограниченному росту величины связи. Скорость обучения зависит от типа связи. Кроме того, различные значения α используются на этапах обучения и тестирования модели. Так, обучение разделено на две стадии с высокими и низкими значениями α . Это позволяет стабилизировать глобальный порядок в картах на финальном этапе обучения. Во время тестирования используются большие значения α для латеральных возбуждающих связей, что способствует надежной синхронизации между нейронами, стимулируемыми одним зрительным объектом.

При обучении (40000 предъявлений различных изображений) в качестве входных изображений использовались объекты (полоски) эллипсоидальной формы,

расположенные на белом фоне. Распределение интенсивности внутри каждого эллипса гауссовское (от черного в центре до серого на краях). Соотношение большого и малого радиусов эллипса принималось $a^2 = 15$, $b^2 = 0,6$ на начальном этапе обучения, а в дальнейшем $a^2 = 45$, $b^2 = 0,45$. Каждый такой объект перекрывал несколько рецептивных полей. Объекты располагались в случайных местах зрительного поля и имели случайную ориентацию. В ходе обучения нейроны постепенно становились чувствительными к определенной ориентации объекта, расположенного в их рецептивном поле. В конце концов в модели возникла глобальная ретинотопическая карта предпочтительных ориентаций, почти идентичная для обоих слоев и сходная с той, которая имеется в зрительной коре (рис. 7). После обучения латеральные связи с низкими значениями весов были удалены. При этом оказалось, что оставшиеся латеральные возбуждающие связи имеют ту же архитектуру, которая была постулирована в модели Ли. Однако следует отметить, что, в отличие от модели Ли, тормозные связи в MAP2 имеют одинаковый вес.

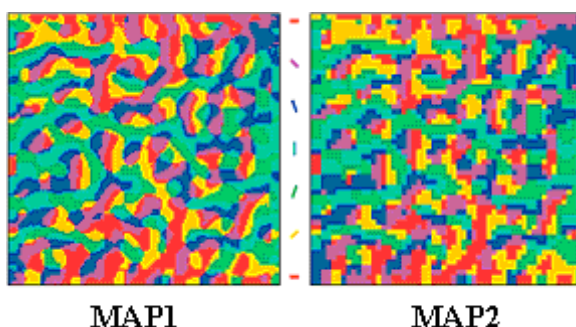


Рис. 7. Распределение ориентационных предпочтений, полученное в результате обучения в картах MAP1 и MAP2. Ориентационные предпочтения обозначены цветами в соответствии с обозначениями углов наклона, показанных в центре рисунка [13].

Работоспособность модели тестировалась на изображениях, содержащих набор из 9 коротких эллипсоидальных полосок (соотношение большого и малого радиусов $a^2 = 3,5$, $b^2 = 1,5$) (рис. 8). Размер полоски достаточно мал, чтобы она помещалась в рецептивном поле одного из нейронов. Полоски располагались в различных местах зрительного поля и имели различную ориентацию. На каждом изображении имелись три более или менее коллинеарных полоски (точность выполнения условия коллинеарности была различной для разных изображений) с небольшими пробелами фона между каждой из пар полосок. Предполагалось, что эти полоски образуют участок контура. Остальные 6 полосок имели случайную ориентацию и играли роль шума.

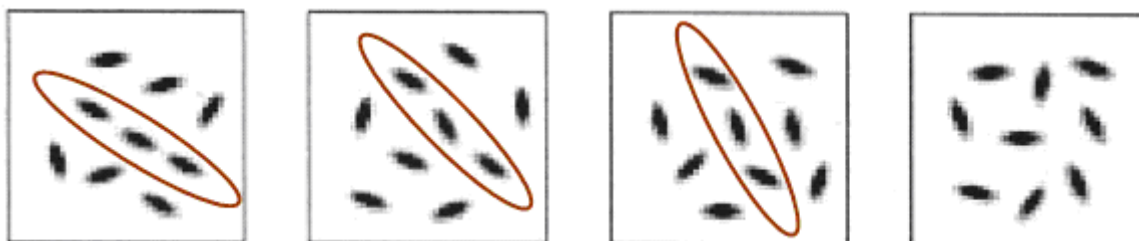


Рис. 8. Примеры изображений, использовавшихся для тестирования модели Чое-Миккулаинена. Полоски, образующие линию, показаны как обведенные красными эллипсами. Рисунки упорядочены слева направо по степени ухудшения коллинеарности полосок, образующих линию. На правом рисунке углы наклона полосок случайны [13].

Целью компьютерных экспериментов было проверить, как влияет точность задания коллинеарности контурных полосок на уровень корреляции стимулируемых ими нейронов. Для этого для каждой полоски вычислялась множественная нейронная активность (МНА), определяемая как число импульсов, генерируемых нейронами, расположенными в области, занимаемой этой полоской. Было показано, что в случае высокой коллинеарности полосок соответствующая им внеклеточная МНА почти идеально синхронна (синфазна). Синхронизация падает по мере ухудшения коллинеарности полосок, образующих контур. В случае случайной ориентации полосок синхронизация отсутствует.

В других экспериментах была обнаружена высокая корреляция между степенью синхронизации МНА и способностью человека к идентификации соответствующих контуров на изображении. Если изображение содержит два контура с разной ориентацией (каждый из которых представлен тремя коллинеарными полосками), то при предъявлении изображения сначала возникает синхронная (синфазная) активность во всех активированных нейронах, но позднее формируется два кластера МНА, работающих в противофазном режиме.

Из экспериментальных данных известно, что качество выполнения сегментации контуров различно в центре поля зрения и на периферии, а также в верхней и нижней зрительных полусферах. Для их объяснения были выдвинуты две гипотезы: эти различия могут быть следствием различий в (1) частоте предъявления различных изображений при обучении, (2) типе варибельности обучающей выборки. Эти гипотезы проверялись двояким образом. В первом эксперименте объекты при обучении предъявлялись в одной области в два раза чаще, чем в другой. Во втором эксперименте углы ориентации полосок в обучающей выборке варьировались равномерно и случайно в интервале $[0^\circ, 10^\circ]$ в одной области и $[0^\circ, 25^\circ]$ в другой. Как можно было ожидать, эксперименты показали, что группировка с помощью синхронизации более надежна там, где имеет место более интенсивное обучение и где разнообразие предъявляемых углов ориентации выше. Последнее особенно значимо для правильной группировки полосок с неидеальной коллинеарностью.

2.3. Модель Ю-Жанга

Еще один подход к сегментации контуров предложен в работе [81]. В значительной мере в нем воспроизведены основные черты модели Ли – трехмерная сеть нейронов, чувствительных к ориентации, сгруппированная в гиперколонки, локальные горизонтальные связи и т.д., но в дополнение к зрительным сигналам, поступающим от рецептивного поля, на нейроны приходят сигналы от областей, моделирующих верхние зоны коры. Эти сигналы используются для выбора определенного контура из набора одновременно заданных контуров, на которые, кроме того, наложен еще дополнительный шум. Другим отличием данной модели от модели Ли является использование в качестве базовых элементов не формальных осцилляторов, а импульсных нейронов, сходных с теми, что фигурировали в модели Чое-Миккулаинена.

Нейроны получают два сигнала: периодическую (с периодом σ) последовательность импульсов от входного зрительного канала и нисходящие сигналы, кодирующие априорную информацию о выбираемом контуре. Априорная информация задается в виде последовательности кодов. Например, если контур состоит из сегментов (полосок) с ориентациями $0, \pi/4, \pi/2, 3\pi/4, \dots$, то нисходящие импульсы будут посылаются параллельно на все нейроны с предпочтительной ориентацией 0 , затем (после временной задержки σ) импульсы будут посылаются на все нейроны с предпочтительной ориентацией $\pi/4$, еще позднее – на все нейроны с ориентацией

$\pi/2$ и т.д. Таким образом, длина нисходящей импульсной последовательности равна числу отрезков в контуре L , а активность нейронной сети имеет период $L\sigma$.

Идея, положенная в основу функционирования модели, – это детекция временного совпадения импульсов при наличии синаптических задержек. Предполагается, что константа временной задержки при обмене сигналами между нейронами равна σ . Вследствие этого генерируемая нейроном активность будет распространяться на соседние клетки, но только в том случае, если они получают подходящий нисходящий сигнал, то есть если контур данного объекта соответствует контуру, заданному априори. Рассогласование между кодирующей априорной последовательностью и реальной ориентацией контурного сегмента контура немедленно прервет "рост" контура.

2.4. Обсуждение

Подведем итоги данного раздела. Модель Ли отражает анатомические и нейрофизиологические данные о функционировании первичных зон зрительной коры и является хорошей основой для дальнейшего развития моделей сегментации контурных изображений. Усовершенствование модели, предложенное Й. Чое и Р. Миккулаином и связанное с автоматическим формированием связей в процессе обучения и функционирования, представляется шагом в перспективном направлении, позволившем включить в модель ряд новых экспериментальных данных. Однако использованные для тестирования модели изображения являются достаточно условными. В работе не приводятся каких-либо результатов сегментирования контуров естественных изображений. Модель Ю-Жанга кажется наиболее спорной, поскольку ее функционирование неустойчиво при наличии шумов и деформаций контуров. Кроме того, представляется правдоподобным, что сегментация контуров может в значительной степени осуществляться автономно первичными зонами зрительной коры без влияния более высоких зон, что противоречит постулатам этой модели. Тем не менее, идея о выборе определенного объекта по информации о его признаках, поступающей сверху вниз, является плодотворной и в несколько других условиях возникает в моделях, которые будут рассмотрены далее (см. раздел 5). Отметим, что класс изображений, с которыми работают существующие контурные модели, является сравнительно узким, а успешных результатов по обработке этими методами достаточно сложных естественных изображений пока что нет.

3. СЕГМЕНТАЦИЯ ПИКСЕЛЬНЫХ ИЗОБРАЖЕНИЙ

В нейросетевых моделях имеется два популярных подхода к решению задачи сегментации изображений путем формирования синфазных нейронных ансамблей, кодирующих разные объекты. В соответствии с первым подходом такие ансамбли формируются одновременно, то есть несколько несинфазных или даже несинхронных между собой ансамблей сосуществуют в одной нейронной сети. При этом синфазность между нейронами одного ансамбля служит меткой, позволяющей реализовать связывание признаков объектов на последующих этапах обработки информации. Согласно второму подходу синфазные ансамбли, кодирующие объекты, возникают последовательно (в детерминированном или случайном порядке), так что в каждый момент времени существует лишь один синхронно работающий ансамбль нейронов с высоким уровнем активности, в то время как активность в других частях нейронной сети временно подавлена до низкого уровня.

3.1. Параллельные процедуры

В случае параллельных процедур задача сегментации изолированных объектов, контрастно выделяющихся из фона, представляется достаточно тривиальной. Она решается путем задания локальных синхронизирующих связей в однослойной сети. Десинхронизация между различными ансамблями наступает за счет действия шумов или за счет дальних десинхронизирующих связей. Конечно, для разработки модели с биологически правдоподобными элементами и с реальными временными характеристиками надо приложить определенные усилия, но принципиальных трудностей на этом пути нет. В связи с этим основное внимание исследователей обращено на работу с реальными изображениями, где объекты могут пересекаться и где нет четкой границы между соседними объектами, а также между объектами и фоном.

Один из наиболее успешных примеров такого рода моделей предложен в работе Опара и Вёргёттера [53]. Схема функционирования модели представлена на рис. 9. Элемент модели описывается дискретной переменной, называемой спином, с конечным числом состояний. Для краткости будем и сами элементы называть спинами. Имеется взаимнооднозначное соответствие между спинами и пикселями изображения. Информация об изображении преобразуется в величины взаимодействия между спинами, зависящие от разности интенсивностей соответствующих пикселей. Динамика модели формулируется в терминах минимизации функции энергии.

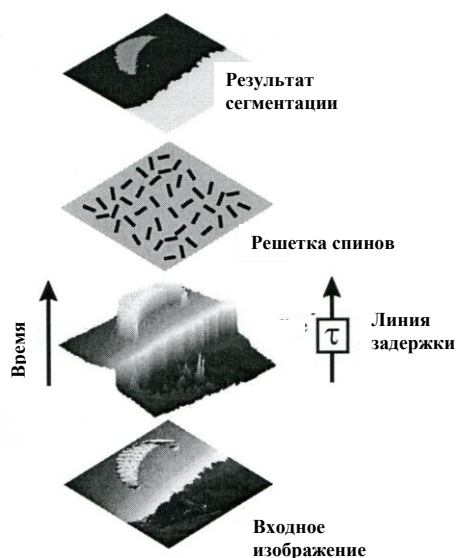


Рис. 9. Схема работы модели Опара-Вёргёттера. Входное изображение подается на линию задержки. Величина задержки зависит от освещенности пикселей [53]. Яркие объекты обрабатываются раньше, чем темные. Временная структура потока информации определяет взаимодействие между спинами на решетке спинов.

Спины расположены в узлах двумерной решетки размером $N = L_x L_y$ с локальными связями между спинами. Каждый спин w может принимать q различных состояний. Номер состояния рассматривается как метка спина. Целью сегментации является такое распределение меток, чтобы разным объектам были присвоены различные метки, а все спины, представляющие один объект, имели одну и ту же метку. Конфигурацией сети называется набор состояний (меток) всех ее спинов $W = \{w_1, \dots, w_N\}$. Глобальная энергия конфигурации W определяется как

$$E(W) = \sum_{i=1}^N (K_i - \sum_{j \in N_i} J_{ij} \delta_{w_i w_j}),$$

где w_i, w_j - метки спинов i и j , соответственно; J_{ij} - сила взаимодействия между спинами i и j ; $\delta_{w_i w_j} = 1$, если $w_i = w_j$, в противном случае $\delta_{w_i w_j} = 0$; N_i - окрестность спина i радиуса Δ ; K_i - величина, описывающая глобальное ингибирование, гарантирующее, что удаленным объектам будут присвоены различные метки:

$$K_i = \frac{\alpha}{N} \sum_{j=1}^N \delta_{w_i w_j},$$

α - положительный параметр.

Значения J_{ij} задаются таким образом, что позитивная сила связи соответствует пикселям со сходными значениями интенсивности, а отрицательная сила связи соответствует пикселям, интенсивности которых значительно различаются:

$$J_{ij} = 1 - \frac{|g_i - g_j|}{\Theta},$$

где g_i, g_j - интенсивности пикселей, Θ - определенным образом усредненное значение разностей интенсивностей.

Традиционные процедуры спиновой динамики с локальными связями имеют низкую скорость адаптации. Кроме того, необходимо бывает принимать специальные меры (типа имитации отжига) для избегания локальных минимумов. Для преодоления этих трудностей авторы предлагают подход, сочетающий алгоритм с локальным взаимодействием и алгоритм кластеризации.

Начальным состоянием модели является случайное распределение меток спинов. Идентичные спины объединяются в кластеры. Формирование кластера имеет стохастическую природу и зависит от температуры T . Если T высока, средняя величина кластера мала. Если T низка, возможный размер кластера увеличивается. Таким образом, температура является параметром, управляющим числом сегментируемых объектов на изображении. После завершения формирования кластеров начинает работать алгоритм модификации кластеров. Для каждого кластера вычисляется функция энергии, зависящая от окружения этого кластера и температуры. Далее метки всех спинов в кластере одновременно устанавливаются так, чтобы удовлетворять гиббсовскому распределению энергии кластера. В каждый момент времени происходит модификация одного кластера при фиксированных метках соседних кластеров.

Этот алгоритм был назван энергетической модификацией кластеров (ЭМК). Для улучшения работы ЭМК авторы вводят зависящие от величины интенсивности пикселей задержки в передаче сигналов от пикселей к спинам (яркие пиксели обрабатываются раньше, чем темные пиксели) [52]. Это гарантирует одновременную активацию спинов, представляющих пиксели со сходными значениями интенсивности. Предполагается, что сила взаимодействия между спинами равна нулю до того момента, когда они получают внешний сигнал. Таким образом, в рассматриваемой системе происходит преобразование структуры интенсивностей во временную структуру входных данных, что в свою очередь оказывает влияние на ход вычисления энергии кластеров и, соответственно, на формируемую структуру спиновых меток.

ЭМК тестировалась на синтетических и реальных изображениях (рис. 10). Было показано, что ее характеристики (быстродействие, качество сегментации, устойчивость к шумам, способность восстанавливать детали изображения) превосходят известные из

работ по искусственному интеллекту алгоритмы, использующие локальную информацию, и алгоритмы, основанные на модификации кластеров. В частности, при низкой температуре для сегментации изображения бывает достаточно всего нескольких десятков итераций ЭМК.

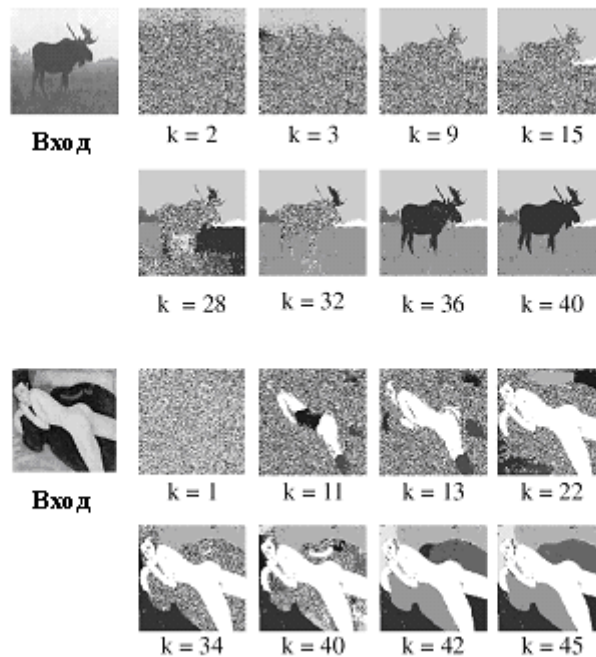


Рис. 10. Результаты сегментации на различных итерациях работы модели Опара-Вёргёттера. Входное изображение 128×128 пикселей [53].

В модели [53] метки сегментов, задаваемые спинами, носят абстрактный характер, не позволяющий непосредственную реализацию модели в виде биологически правдоподобной нейронной сети. В работе Гричука с соавторами [91] предложена модель, реализующая сегментацию в сети из осцилляторных элементов. Сеть состоит из осцилляторов предельного цикла, локализованных в узлах двумерной решетки. Каждому пикселю соответствует один осциллятор решетки. Уравнения динамики одного слоя записываются в комплексных переменных как

$$\frac{du_i}{dt} = f(u_i, \rho_i, \omega_i) + \sum_{j=1}^N w_{ji}(u_j - u_i), \quad i = 1, \dots, N,$$

где u_i - комплексная переменная, N - число осцилляторов в слое, w_{ji} - весовые коэффициенты/ Функция f определяет динамику изолированного осциллятора предельного цикла с амплитудой ρ_i и собственной частотой ω_i . Веса w_{ji} заданы в виде нелинейных функций, зависящих от ρ_i и ρ_j и убывающих по мере увеличения расстояния между осцилляторами. Амплитуды ρ_i и частоты ω_i задаются как монотонные функции от яркости пикселей.

Сегментация изображения осуществляется в два этапа. Сначала в соответствии с величинами яркости настраиваются параметры сети. Затем запускается динамика сети, приводящая к появлению синфазных ансамблей осцилляторов, представляющих сегменты изображения.

Как показано в работах [39, 90, 94] однослойная сеть такой конструкции позволяет сегментировать изображение в том случае, когда различные сегменты имеют

существенно отличные значения яркостей или цвета. Для обобщения сегментации на случай, когда сегменты имеют близкие оптические характеристики, использована двухслойная архитектура с независимыми слоями, отличающимися только начальными состояниями. Начальное распределение фаз в слоях задавалось пропорционально соответственно x и y координатам осцилляторов. Благодаря такому выбору начальных состояний после релаксации в состояние с синфазными сегментами имел место сдвиг фаз между синфазно синхронизованными ансамблями осцилляторов в одном или обоих слоях. Единственным ограничением было отсутствие пересечения сегментируемых объектов хотя бы по одной координате.

Пример сегментации изображения описанным методом приведен на рис. 11. Хотя результат сегментации показан в виде последовательных изображений объектов, на самом деле имеет место сдвиг по фазе между соответствующими ансамблями осцилляторов.

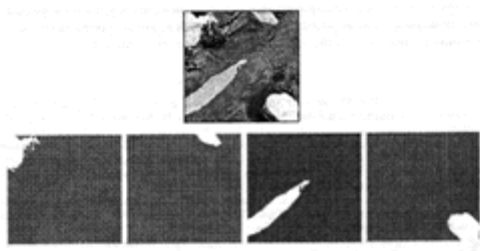


Рис. 11. Сегментация изображения с помощью модели [91]. Сверху исходное изображение. Внизу изображения четырех сегментированных объектов.

3.2. Последовательные процедуры

В отличие от модели раздела 3.1, в которой метки, присваиваемые сегментам изображения, носят формальный характер, в данном разделе мы рассмотрим более традиционный для динамических нейронных сетей подход к сегментации, основанный на колебаниях и синхронизации нейронной активности. Последовательная сегментация в данном случае означает, что в какой-то период времени в модели возникает синхронно работающий ансамбль осцилляторов с высокой активностью, геометрия которого совпадает с геометрией одного из объектов на входном изображении. В следующий период времени в сети возникает другой синхронный ансамбль, соответствующий другому объекту, и т.д. Последовательность, в которой осуществляется сегментация, обычно имеет стохастический характер с большим приоритетом для объектов с определенными геометрическими или спектральными свойствами. Например, объекты выбираются в порядке убывания их площади, причем совсем маленькие объекты игнорируются как шум.

Для реализации конкуренции между нейронными ансамблями, соответствующими сегментируемым объектам, имеется две возможности. Первая – это использовать дальнедействующие тормозные связи. Однако при большом числе элементов в сети это приводит к слишком большому числу связей. Другая возможность – использовать специальный центральный элемент, играющий роль глобального тормозного элемента (*global inhibitor*), связанного со всеми элементами сети. В этом случае число связей в сети имеет тот же порядок, что и число элементов. Данный принцип реализован в модели Ванга и Термана, называемой LEGION (*Locally Excitable, Globally Inhibitory Oscillator Network*) [9 - 10, 74 - 76]. Первоначальный вариант модели был рассчитан на работу с изолированными черными объектами на белом фоне. Позднее авторы вносили в базовую модель различные усовершенствования, позволившие сегментировать достаточно сложные полутоновые изображения.

В сети LEGION осцилляторы расположены в узлах плоской решетки, так что каждый осциллятор ассоциирован с одним пикселем на входном изображении (рис. 12). Между осцилляторами имеются локальные возбуждающие связи, сила которых убывает по мере увеличения расстояния между осцилляторами. С помощью этих связей формируются ансамбли синхронно работающих осцилляторов. В сети имеется также глобальный тормозный элемент (ГТЭ), предназначенный для реализации конкуренции между ансамблями осцилляторов и препятствующий глобальной синхронизации всех элементов сети.

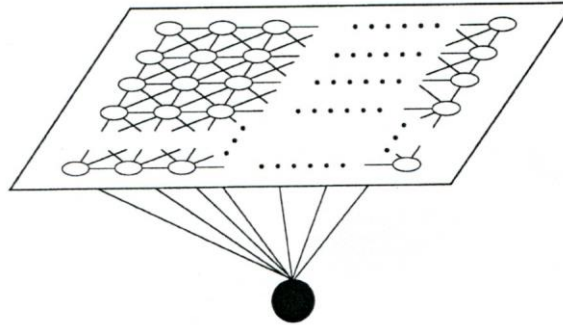


Рис. 12. Архитектура связей в модели LEGION. Полыми кружками обозначены осцилляторы, черным кружком обозначен глобальный тормозный нейрон [9].

Элементом модели является модифицированный релаксационный осциллятор Ван-дер-Поля, динамика которого описывается следующими уравнениями (для упрощения обозначений индекс, показывающий номер осциллятора, опущен всюду, где это не приводит к недоразумениям)

$$\frac{dx}{dt} = 3x - x^3 + 2 - y + R + S + \rho,$$

$$\frac{dy}{dt} = \varepsilon(\alpha(1 + \tanh(\beta x)) - y).$$

Здесь R – величина, зависящая от внешнего входа (сигнала от пикселя), модулируемого сигналами от соседних осцилляторов, S – интегральный сигнал от соседних осцилляторов и ГТЭ, ρ – гауссовский шум, α, β – положительные параметры, $0 < \varepsilon \ll 1$ – малый параметр, обеспечивающий генерацию релаксационных колебаний.

Активность ГТЭ определяется как

$$\frac{dz}{dt} = \phi(\sigma_\infty - z),$$

где $\sigma_\infty = 1$, если величина x_i превосходит порог θ_z для хотя бы одного осциллятора, в противном случае $\sigma_\infty = 0$; $\phi \gg \varepsilon$ – параметр.

Рассмотрим изолированный осциллятор без шума при фиксированной величине члена R . Заметим, что x -нульклина – кубическая функция, а y -нульклина – сигмоидальная функция (рис. 13).

Если $R > 0$, обе нульклина пересекаются в единственной неустойчивой точке, расположенной в средней части кубической кривой. Это приводит к появлению устойчивого предельного цикла, в котором переменная x принимает то высокие (*активная фаза*), то низкие (*фаза молчания*) значения. Параметры осциллятора выбраны таким образом, что в течение каждой из фаз величина x меняется медленно, а при переходе от одной фазы к другой изменение происходит практически скачком. Если $R \leq 0$, нульклина пересекаются на левой ветви кубической кривой в двух точках,

одна из которых устойчивая. В этом случае периодические колебания становятся невозможными. Осциллятор стабилизирует свою активность при низких значениях x .

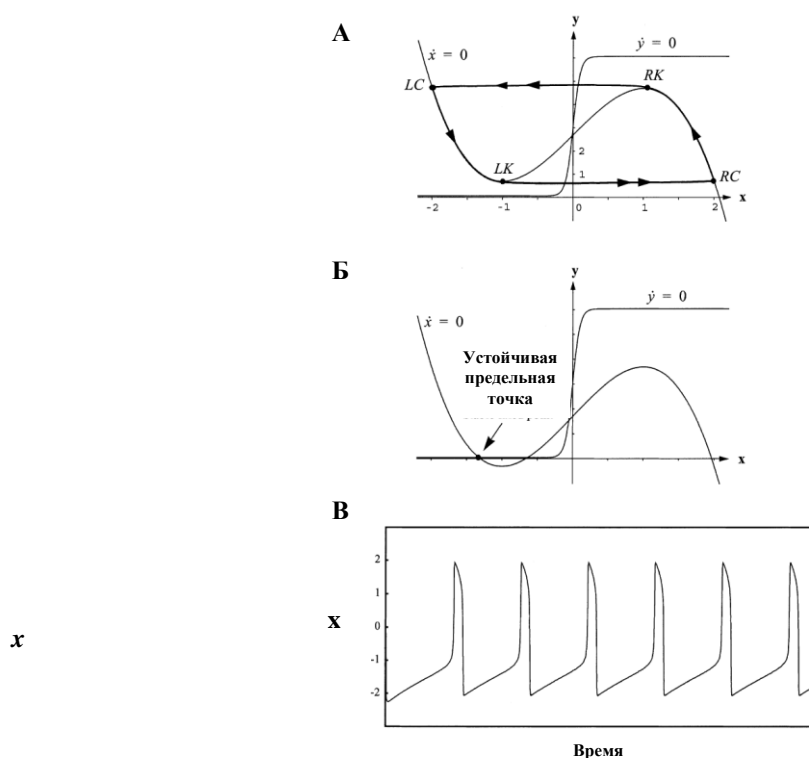


Рис. 13. Фазовый портрет и траектория осциллятора Ванга-Термана [9]. **А.** Осцилляторный режим. Кривая LC-LK-RC-RK показывает предельный цикл осциллятора. **Б.** Стационарный режим осциллятора. **В.** Траектория переменной x .

Синфазные ансамбли формируются в сети благодаря локальным связям между осцилляторами. Введенный в модель ГТЭ играет роль устройства, определяющего моменты времени, когда ансамбль осцилляторов становится активным или переходит в фазу молчания. Ансамбли конкурируют за влияние, оказываемое на ГТЭ, в то же время стараясь избежать обратного тормозного воздействия. Ансамбль большего размера и с более быстро растущей активностью имеет больший шанс "победить" в этой конкуренции. В течение определенного временного интервала активность этого ансамбля остается высокой, а активность других ансамблей подавляется ГТЭ до уровня молчания. Однако такое состояние системы не длится долго. Благодаря действию ГТЭ, активность во всей сети падает до низкого уровня, вследствие чего активность ГТЭ падает до 0, и далее процесс конкуренции начинается заново с тем, чтобы привести к "победе" другой ансамбль осцилляторов. Параметры взаимодействия осцилляторов с ГТЭ подобраны следующим образом: тот ансамбль, который был активен, подвергается со стороны ГТЭ настолько сильному торможению, что не может участвовать в конкуренции в течение времени, достаточного для последовательной активации остальных ансамблей. Это позволяет системе последовательно активировать различные ансамбли осцилляторов. Стохастический характер порядка выбора объектов обусловлен наличием шума.

Какие именно ансамбли формируются в сети, зависит от задания членов R и S , поэтому основные усилия разработчиков модели были направлены на такое задание этих величин, чтобы адаптировать процесс формирования ансамблей к сегментации реальных изображений в присутствии шума. Численные эксперименты показали, что константные значения R и S (зависящие исключительно от входного изображения) не

дают качественной и устойчивой сегментации. В связи с этим как для внешних сигналов, так и для величин связей были использованы адаптируемые переменные. Ключевым в этой связи явилось понятие "лидера", введенное в работе [76]. Лидер – это осциллятор, расположение которого в сети соответствует центру достаточно большой гомогенной области изображения. Фрагмент изображения, порожденный шумом, не содержит лидера. Лидер получает значительное возбуждение от соседних осцилляторов, переходит в активное состояние и становится ядром, из которого начинается формирование связанного ансамбля активных осцилляторов. Активное состояние распространяется по осцилляторам от лидера к границам сегментируемой области, и там распространение останавливается благодаря подобранным величинам связей. Член R ответственен за идентификацию лидеров, а член S определяет структуру величин связей.

Развитие модели претерпело несколько модификаций, отличающихся тем, как задавались значения R и S . Ниже приводится описание последнего из опубликованных авторами вариантов, модель Чена-Ванга [9].

Для определения осцилляторов-лидеров вводится бинарная переменная p :

$$p = H(T_p - \sigma_{N_p}),$$

где H - функция Хевисайда

$$H(v) = \begin{cases} 1, & v \geq 0, \\ 0, & v < 0, \end{cases}$$

T_p - порог, σ_{N_p} - стандартное отклонение, описывающее уровень гомогенности в квадратной окрестности N_p размера $(2\rho + 1) \times (2\rho + 1)$ вокруг данного пикселя (T_p и ρ - фиксированные параметры).

Переменная R определяется как

$$R = Ip,$$

где I - интенсивность пикселя.

Переменная S определяется как

$$S = WH \left(\sum_{k \in N_1} H(x_k) - 0.5 \right) - W_z H(z - \theta_z),$$

где H – функция Хевисайда; W – динамический вес, задающий силу влияния на данный осциллятор его ближайшей окрестности N_1 ; W_z - сила тормозного воздействия на осциллятор; θ_z - порог (W_z и θ_z - фиксированные параметры).

Переменная W бинарная: $W = 1$, если в текущий момент времени пиксель идентифицируется как лежащий внутри сегментируемой области, $W = 0$, если пиксель идентифицируется как принадлежащий границе. Значение W вычисляется в два этапа. Сначала для каждого пикселя определяются 4 специальные окрестности (одна из них фиксированная, а три других динамические), которые зависят от распределения активных и молчащих осцилляторов в окрестности данного осциллятора. Затем для этих окрестностей вычисляются средние и стандартные отклонения интенсивностей. Значение W вычисляется как функция этих величин. Дополнительно те же величины используются для вычисления значений непрерывной переменной $u \in (0,1)$, которая показывает, принадлежит ли данный пиксель внутренней части сегментируемой области или расположен вблизи границы. В свою очередь, u влияет на формирование окрестностей пикселя на следующем шаге итерации.

При фиксированных значениях параметров число сегментов, которые способна сформировать сеть LEGION, не велико. Более или менее надежной работы модели можно добиться при числе сегментируемых объектов, не превышающем 5-7. На самом деле, это число определяется отношением периода колебания в сети (время просмотра всех объектов) ко времени активной фазы для одного объекта. Для увеличения числа сегментируемых объектов авторы рекомендуют применение иерархической модели, состоящей из нескольких сетей LEGION, в которой сначала выполняется грубая сегментация, которая затем делается более детальной путем дальнейшего дробления составных объектов. Предложена также алгоритмическая реализация сети LEGION, в которой ограничения на число сегментируемых объектов сняты.

Многочисленные эксперименты по сегментации изображений с помощью сети LEGION включали в себя искусственные изображения с высоким уровнем шумов, магниторезонансные томограммы мозга, аэрофотоснимки гидрографических объектов и др. Полученные результаты сегментации сравнивались с традиционными методами сегментации, основанными на адаптивных методах и нелинейном сглаживании [7, 77]. Было продемонстрировано, что LEGION не уступает в качестве сегментации этим методам, а в ряде случаев дает лучшие результаты. Примеры сегментации с помощью сети LEGION приведены на рис. 14 - 15.

Модель сегментации, предложенная Урсино с соавторами в работах [71 - 72], во многих отношениях сходна с сетью LEGION. В ней также используются принципы локального возбуждающего и глобального тормозного действия на элементы системы. Отличие состоит в том, что в данной модели сегментация объектов и группировка пикселей внутри каждого объекта рассматриваются как относительно независимые операции. В связи с этим, модель соединяет в себе традиционные алгоритмы компьютерного зрения для выполнения сегментации объектов и методы синхронизации колебаний для формирования синфазных нейронных ансамблей, представляющих объекты в сети.

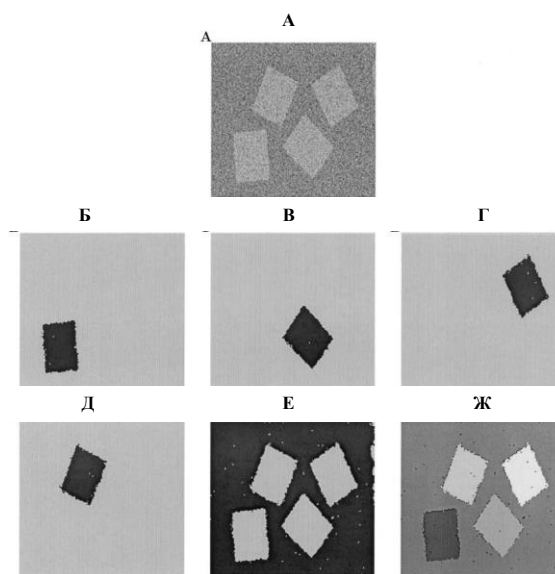


Рис. 14. Сегментация искусственного изображения с добавленным шумов [9]. А. Исходное изображение 230×240 пикселей. Б – Е Последовательное выделение сегментов (включая фон). Ж. Результат сегментации. Сегменты окрашены в разные оттенки серого.

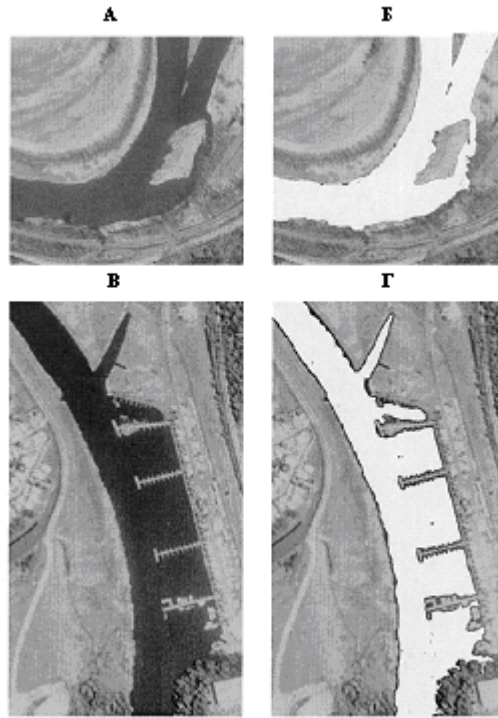


Рис. 15. Сегментация аэрофотоснимков гидрографических объектов моделью LEGION [9]. **А.** Изображение 640×640 пикселей. **Б.** Результат сегментации изображения А. **В.** Изображение 378×670 пикселей. **Г.** Результат сегментации изображения В.

Модель имеет два слоя элементов. Первый слой формируется из неосцилляторных элементов и предназначен для выделения контуров объектов. Второй слой реализует интеграцию пикселей на основе временной корреляции. В качестве второго слоя используется упрощенный вариант сети LEGION с постоянными величинами связей между элементами. Информация о контурах, получаемая в первом слое, используется для подавления взаимодействия между ансамблями осцилляторов, представляющих различные объекты с тем, чтобы не допустить глобальную синхронизацию всей сети. Для улучшения и ускорения синхронизации используются как возбуждающие, так и тормозные связи между соседними осцилляторами.

Выходной сигнал ГТЭ описывается выражением

$$z = A \left[\text{sign} \left(\sum_{i=1}^N x_i - \theta \right) + 1 \right],$$

где x_i - состояние возбуждающего элемента i -го осциллятора, θ - порог, A - сила глобального торможения.

Контурные объекты, вычисленные в первом слое, используются как сильные тормозные сигналы, идущие вертикально снизу вверх на осцилляторы второго слоя. Эти сигналы подавляют осцилляторную активность "контурных" осцилляторов, создавая таким образом изолированные (не взаимодействующие друг с другом) ансамбли осцилляторов, соответствующие отдельным объектам на изображении.

В работе [71] было показано, что результаты сегментации существенно зависят от параметров ГТЭ A и θ . Например, высокие значения обеих величин A и θ могут привести к тотальному подавлению активности маленьких по размеру ансамблей осцилляторов, в то время как комбинация высокого значения θ с низким значением A позволяет маленьким ансамблям активироваться одновременно, формируя диффузно

распределенный объект. Низкие значения A и θ должны использоваться, если мелкие детали изображения представляют важность, однако в этом случае возникает опасность, что некоторые маленькие ансамбли будут активироваться одновременно с каким-то большим ансамблем.

Используя традиционные алгоритмы выделения контуров, основанные на гауссовских фильтрах, авторы показали, что модель успешно работает на искусственных изображениях с высококонтрастными объектами.

3.3. Обсуждение

Вопрос о том, какие механизмы сегментации, последовательные или параллельные, используются мозгом, на экспериментальном уровне не решен. В связи с этим оба класса моделей имеют право на существование. Вполне возможно, что оба механизма задействованы мозгом. Параллельный соответствует предвнимательной обработке информации, а последовательный имеет место при наличии внимания [3, 33]. Если эта гипотеза верна, то параллельные методы должны быть менее точными, но более быстрыми. К сожалению, на уровне имеющихся модельных разработок проверить эту гипотезу не представляется возможным, поскольку существующие модели оперируют разнородными формальными элементами, конструкция которых далека от биологических нейронов.

Модели Опара-Вёргёттера и Ванга с соавторами являются в настоящее время наиболее продвинутыми и успешными в анализе реальных изображений по сравнению с другими моделями в своем классе методов (параллельных или последовательных). Очевидное их отличие состоит в том, что модель Опара-Вёргёттера носит абстрактный характер, для ее нейронной реализации потребуются значительные дополнительные усилия, в то время как LEGION стремится, по возможности, соответствовать нейробиологическим механизмам работы мозга. С точки зрения алгоритмической эффективности основным недостатком модели Опара-Вёргёттера является то, что она рассчитана на класс изображений, в которых каждый индивидуальный объект содержит пиксели со сходными оптическими характеристиками. Это далеко не всегда имеет место в реальности.

Модель Гричука с соавторами [91] имеет существенное ограничение на взаимное расположение объектов. Кроме того, не ясна разрешающая способность модели при близком расположении объектов в пространстве. Предположение о зависимости начальных фаз осцилляторов от координат осцилляторов не имеет экспериментального подтверждения.

Параллельные методы сегментации в традиционном нейронно-сетевом оформлении предлагались в работах [63, 73]. Изложение этих работ имеется в обзоре [92]. К сожалению, несмотря на то, что модели Кёнига с соавторами были приближены к биологической реальности и соответствовали ряду экспериментальных данных, авторы ограничились лишь применением их к сегментации простых искусственных изображений. Продолжения эти работы не получили, по всей видимости, из-за громоздкости моделей.

Модель Чена-Ванга была получена в результате многолетних усилий по усовершенствованию принципов функционирования сети LEGION. В ранних вариантах модели взаимодействие между двумя соседними осцилляторами зависело только от значений интенсивности соответствующих пикселей [76]. Однако оказалось, что качество сегментации существенно повышается при непарном локальном взаимодействии осцилляторов [9]. Плата за это – высокая сложность модели и отсутствие биологической аргументации в поддержку ряда используемых механизмов. В частности, модификация силы связей, по-видимому, предполагает наличие быстро модифицирующихся синапсов, существование которых в мозгу до сих пор остается

дискуссионным. Проверка соответствия модели временным характеристикам реальных процессов сегментации также не представляется возможной из-за выбранной конструкции осциллятора.

Модификация сети LEGION, предложенная Урсино с соавторами, как будто бы сильно упрощает конструкцию Чена-Ванга. На самом деле, эта модель просто переносит цент тяжести решаемой задачи на процедуру выделения контуров, так что качество работы всей системы радикально зависит от качества реализации этой процедуры. К сожалению, известные методы выделения контуров не являются идеальными. На реальных изображениях получаемые с их помощью контуры могут не быть замкнутыми и иметь неравномерную толщину. Кроме того, они могут порождать ложные контурные фрагменты по всему изображению. Для преодоления этих трудностей требуются достаточно сложные процедуры, что существенно уменьшает привлекательность данного подхода. Однако этот подход может оказаться весьма перспективным, если процедура улучшения и фильтрации "сырых" контуров будет реализована не методами искусственного интеллекта, а на осцилляторных нейронных сетях с использованием тех же механизмов, что и в существующих моделях. Такая задача кажется вполне разрешимой.

4. СЕГМЕНТАЦИЯ И КЛАСТЕРИЗАЦИЯ

Кластеризация – это процедура, позволяющая сгруппировать дискретное множество точек P в многомерном пространстве S в компактные классы (кластеры) на основе определенной в S метрики (расстояния между точками). Термин кластеризация употребляется как синоним классификации без учителя, поскольку предполагается, что задача кластеризации должна решаться без какой-либо априорной информации о принадлежности точек к классам. Задача сегментации может быть сведена к задаче кластеризации следующим образом. Поставим в соответствие каждому пикселю вектор его оптических характеристик (признаков). Тогда, после выбора подходящей метрики в пространстве S таких векторов, проведение их кластеризации даст некоторое решение задачи сегментации. Пиксели, расположенные близко друг к другу и имеющие сходные характеристики, будут иметь приоритет по включению в один и тот же кластер и, следовательно, в один и тот же сегмент изображения. Успешность сегментации будет зависеть как от введенной метрики, так и от алгоритма кластеризации.

Какова связь этих общих соображений с осцилляторными нейронными сетями? Дело в том, что синхронизация колебаний дает удобный способ реализации кластеров. Поставим в соответствие каждой точке P осциллятор и введем в сети из этих осцилляторов взаимодействие таким образом, чтобы осцилляторы, соответствующие близко расположенным точкам в S , были связаны синхронизирующими связями, а осцилляторы, соответствующие удаленным точкам из S , были связаны десинхронизирующими связями. Тогда динамика активности в сети автоматически приведет к формированию синхронных ансамблей осцилляторов, представляющих компактные кластеры точек в S .

4.1. Модель Роума-Фригуи

Эти идеи были применены к сегментации изображений в работе Роума и Фригуи [59]. Авторы ее не слишком заботились о биологической правдоподобности модели. Их целью было выяснить, насколько осцилляторные нейронные сети могут быть эффективными при сегментации изображений. В качестве элементов модели были выбраны так называемые фазово-импульсные элементы. Это было обусловлено тем, что сети из таких элементов (pulse-coupled neural networks - PCNN) являются популярным объектом нейросетевого моделирования и уже зарекомендовали себя как эффективное средство при решении задач по анализу изображений [46].

Фазово-импульсный осциллятор описывается уравнением [49]

$$x = f(\phi),$$

где x – состояние осциллятора, $\phi \in [0,1]$ – его текущая фаза колебаний, $f : [0,1] \rightarrow [0,1]$ – гладкая монотонно-возрастающая функция, $f(0) = 0$ и $f(1) = 1$. Осциллятор генерирует импульс в тот момент, когда достигается состояние $x = 1$. Немедленно после этого фаза обнуляется, $\phi = 0$. При отсутствии входных импульсов фаза меняется как

$$\phi = \omega t,$$

где t – время, прошедшее после последнего импульса, ω – собственная частота осциллятора.

Если в момент t на вход осциллятора приходит импульс с амплитудой ε , фаза осциллятора скачком меняется с $\phi(t)$ на

$$\phi(t^+) = f^{-1}(B(x(t) + \varepsilon)),$$

где f^{-1} – обратная функция и

$$B(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 0, & x < 0, \\ 1, & x > 1. \end{cases}$$

В соответствии с этими формулами, в момент t состояние осциллятора меняется с $x(t)$ на

$$x(t^+) = B(x(t) + \varepsilon).$$

Функция B используется для ограничения интервала изменения x единичным отрезком $[0,1]$.

В численных экспериментах использовалась функция f следующего вида:

$$f(\phi) = \frac{1}{b} \ln[1 + \phi(e^b - 1)],$$

где b – постоянный параметр, задающий меру выпуклости функции f . Все собственные частоты одинаковы. Величина амплитуды импульса ε не постоянна, а меняется в зависимости от текущих фаз пары взаимодействующих осцилляторов (см. ниже).

Опишем теперь, как сеть из фазово-импульсных осцилляторов может использоваться для сегментации. Рассмотрим множество точек $P_i \in S$ ($i = 1, \dots, N$).

Используя координаты точек и метрику в пространстве S , можно вычислить так называемую реляционную матрицу $R = \{r_{ij}\}$, где r_{ij} удовлетворяет соотношениям

$$r_{ij} \in [0,1], \quad r_{ii} = 0, \quad r_{ij} = r_{ji}$$

и характеризует относительное несходство между точками P_i и P_j .

После того как каждой точке P_i поставлен в соответствие фазово-импульсный осциллятор, проблема кластеризации сводится к разбиению множества из N осцилляторов на некоторое число ансамблей синфазных осцилляторов. Заметим, что требование отсутствия синфазности между ансамблями осцилляторов не препятствует тому, чтобы эти ансамбли работали синхронно с некоторым сдвигом фаз относительно друг друга. Число ансамблей считается априорно неизвестным. Оно зависит от параметров алгоритма кластеризации.

Для формирования требуемых ансамблей осцилляторов авторами предложен следующий метод задания амплитуд импульсов:

$$\varepsilon_{ij} = \begin{cases} C_E \left[1 - \left(\frac{r_{ij}}{\delta} \right)^2 \right], & r_{ij} \leq \delta, \\ C_I \left[\left(\frac{1-r_{ij}}{1-\delta} \right)^2 - 1 \right], & r_{ij} > \delta. \end{cases}$$

В этой формуле C_E - максимальная величина возбуждающей связи, C_I - максимальная величина тормозной связи, $0 < \delta < 1$ - параметр, определяющий степень сходства между осцилляторами, при которой они должны быть включены в один ансамбль. Если $r_{ij} \leq \delta$, взаимодействие между осцилляторами возбуждающее, в противном случае – тормозное. Чем меньше δ , тем большее число сегментов будет выделено на изображении.

При формировании в сети ансамбля G синфазных осцилляторов необходимо принять специальные меры к тому, чтобы уже достигнутая синфазность сохранялась в будущем. Для этой цели импульсы, приходящие на осцилляторы одного ансамбля, должны быть одинаковой амплитуды. Это достигается пересчетом значений расстояний r_{ij} ($j \in G$) по формуле

$$r_{ij} = \frac{1}{g} \sum_{j \in G} r_{ij},$$

где g - число осцилляторов в G (эта процедура называется усредняющим алгоритмом связывания).

В начальном состоянии фазы осцилляторов имеют случайные значения. В ходе работы модели осцилляторы начинают объединяться в ансамбли с одновременной импульсацией. Осцилляторы, достигнув своей активности, остаются таковыми благодаря идентичным входным сигналам. Постепенно уже сформировавшиеся ансамбли вовлекают в себя другие осцилляторы. Более того, сами ансамбли могут сливаться, образуя ансамбли большего размера. В конце концов в системе формируются устойчивые ансамбли осцилляторов, генерирующие импульсы с фиксированным фазовым сдвигом между ансамблями.

Описанная выше модель была бы вполне работоспособна, если бы не одна проблема – огромный размер реляционной матрицы в случае работы с реальными изображениями. К счастью, имеется эффективный способ радикально уменьшить объем вычислений расстояний. Для этого вводится понятие *прототипа*. Прототипами называются точки из множества $B = \{b_m\} (m = 1, \dots, M)$, $b_m \in S$, $M \ll N$. Первоначально прототипы случайно распределяются в S . Как только некоторый осциллятор генерирует импульс, определяется ближайший к нему прототип b_m , после чего расстояние от данного осциллятора до ансамбля A заменяется на расстояние от b_m до A . Если осциллятор включается в ансамбль B , то b_m берет на себя роль представителя этого ансамбля, то есть, b_m перемещается в центр B и используется в последующих вычислениях расстояний вместо всех индивидуальных элементов B .

Примеры кластеризации с помощью модели Роума-Фригуи приведены на рис. 16. На рисунке показана динамика кластеризации смеси из шести двумерных гауссовских распределений в случае отсутствия шумов (верхний ряд) и при наличии шумов (нижний ряд). На итерации, показанной на рис. 16А, успевают сформироваться четыре

маленькие группы. Их центры показаны знаком +. Эллипсы показывают окрестности с возбуждающим взаимодействием. Отдельные точки показывают осцилляторы, которые еще не вошли в синхронизацию. Некоторые точки попадают внутрь эллипсов, что означает, что они близки к синхронизации с данной группой, но это потребует большего времени. На рис. 16Б большинство осцилляторов уже входят в группы синхронизованных осцилляторов. На рис. 16В уже практически все осцилляторы вошли в синхронизацию в своих группах. На рис. 16Г - Е видно, что "шумовые" точки или образуют очень маленькие кластеры или вообще не входят в синхронизацию.

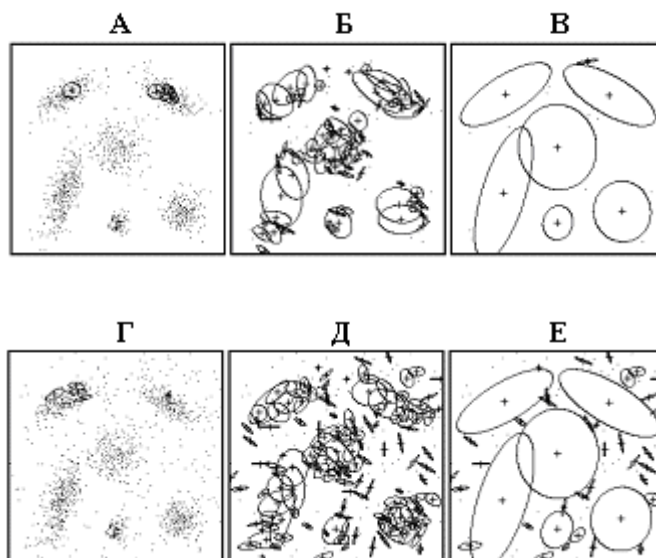


Рис. 16. Кластеризация смеси двумерных гауссовских распределений с помощью модели Роума-Фригуи [59]. Центры сформировавшихся групп показаны знаком +. А - В. Смесь без шума. Результаты кластеризации: после 5 итераций – А, после 500 итераций – Б, после 1200 итераций – В. Г - Е. Смесь с шумом. Результаты кластеризации: после 5 итераций - Г, после 500 итераций – Д, после 1700 итераций – Е.

Модель применялась к сегментации цветных изображений. Каждый пиксель изображения характеризовался восемью признаками. Помимо двух значений координат пикселя в число признаков были включены интенсивности по трем цветам и три текстурные характеристики окрестности пикселя [8]. На рис. 17 показаны результаты сегментации для изображений, взятых из базы Corel Image Database. При $\delta = 0,2$ крупные области на изображении разбиваются на небольшие подобласти, не слишком сильно отличающиеся по своим оптическим характеристикам. При $\delta = 0,3$ формируются крупные области, соответствующие основным объектам на изображении.

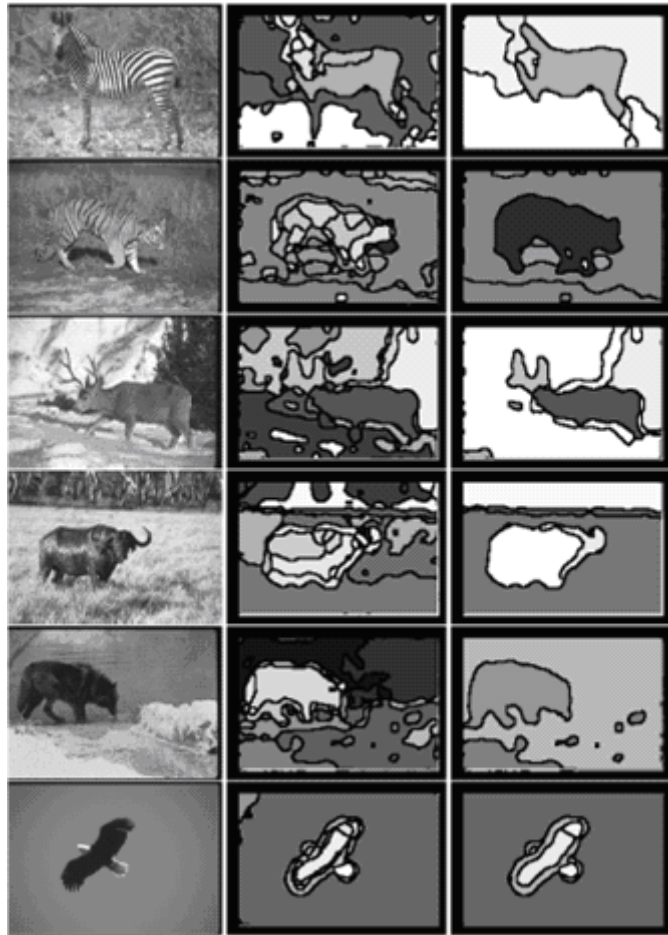


Рис. 17. Сегментации изображений моделью Роума-Фригуи. Левая колонка – исходные изображения. Средняя колонка – сегментация при $\delta = 0,2$. Правая колонка – сегментация при $\delta = 0,3$ [59].

4.2. Модель Жао

В работах Жао с соавторами [83-85] разработан сходный подход к сегментации, но на базе более абстрактных элементов, в качестве которых использованы логистические отображения

$$x(t+1) = f(x(t)) = rx(t)(1-x(t)),$$

где $x(t)$ меняется в интервале $(0,1)$. Известно, что логистическое отображение может породить различные типы динамики в зависимости от значения параметра r . В частности, для большинства значений r между 3,57 (приблизительно) и 4 логистическое отображение порождает хаотическую динамику [70]. Сети из логистических отображений, называемые глобально взаимодействующими отображениями (Globally Coupled Maps - GCM), уже использовались ранее в задачах кластеризации [31]. Модель Жао является обобщением этих сетей с дополнительной адаптацией входных данных.

Рассматриваемая модель представляет собой N взаимодействующих хаотических логистических отображений, каждое из которых соответствует некоторому пикселю на изображении. Динамика сети описывается следующими уравнениями:

$$x_i(t+1) = (1-\varepsilon)f(x_i(t)) + \frac{\varepsilon}{M_i(t)} \sum_{j=1}^N z_{ij}(t)f(x_j(t)), \quad i = 1, \dots, N,$$

где $z_{ij}(t)$ - силы взаимодействия, $M_i(t)$ - число пикселей в локальной окрестности $\Delta_i(t)$ пикселя i , ε - параметр ($0 < \varepsilon < 1$).

Величины $z_{ij}(t)$ зависят от признаков пикселей i и j ,

$$z_{ij}(t+1) = \beta z_{ij}(t) + (1 - \beta)H(e^{-\alpha|c_i(t) - c_j(t)|} - \theta),$$

где $c_i(t) = (c_{i1}(t), \dots, c_{iK}(t))$ - вектор признаков i -го пикселя, $|| \cdot ||$ - евклидова норма, α, β - параметры ($\alpha > 0, 0 < \beta < 1$), H - функция Хевисайда, θ - порог. Параметры α and θ задают уровень различий между сегментами изображения: бóльшим значениям этих параметров соответствует большее число кластеров.

Особенностью модели является адаптируемость векторов $c_i(t)$ в соответствии с "силой", действующей на пиксель i со стороны пикселей из $\Delta_i(t)$. Эта сила велика/мала в зависимости от того, имеют ли пиксели сходные/несходные признаки. При работе модели координаты векторов постепенно сдвигаются в направлении действия равнодействующей силы со стороны $\Delta_i(t)$. Благодаря подходящему выбору процедуры адаптации, точки, отнесенные к одному кластеру, группируются в нем более тесно, а расстояние между кластерами возрастает (рис. 18).

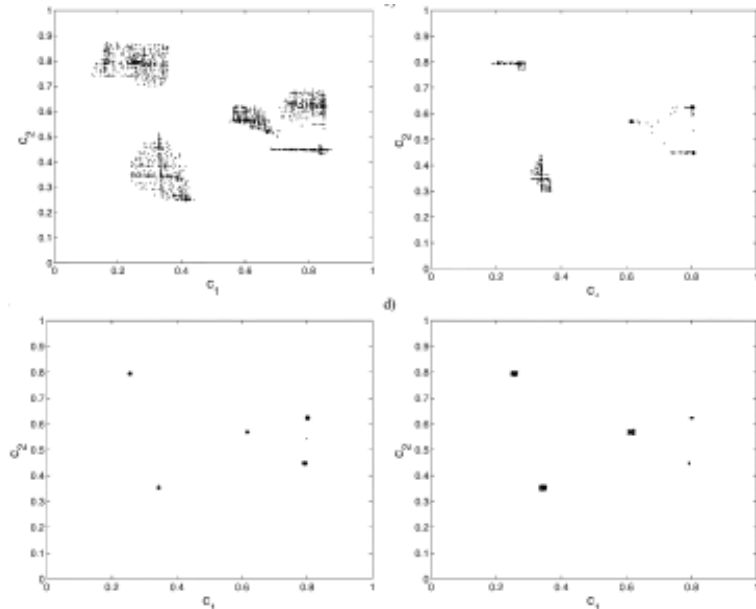


Рис. 18. Формирование кластеров в результате движения изображающих точек. На рисунках показаны состояния шести кластеров в моменты времени 1, 20, 30, 55. Видно, что размеры кластеров постепенно уменьшаются [83].

Процесс кластеризации описывается следующим образом. В исходном состоянии силы взаимодействия ε и $z_{ij}(0)$ устанавливаются на достаточно высокий уровень, чтобы синхронизовать всю сеть. При подаче на вход модели изображения каждая переменная $z_{ij}(t)$ стремится к значению, даваемому функцией Хевисайда. Эта величина определяет окрестность пикселя $\Delta_i(t)$, пиксели которой считаются сходными по своим признакам с пикселем i . Другими словами, взаимодействие между осцилляторами i и j становится максимальным, когда $z_{ij}(t)$ приближается к 1, а при $z_{ij}(t)$, приближающемся к нулю, взаимодействие исчезает. Постепенно все осцилляторы, стимулируемые сходными входными сигналами, формируют синфазный хаотический ансамбль. Осцилляторы, стимулируемые пикселями со значительно различающимися

признаками, будут отнесены к разным ансамблям. Преимущества использования хаотической динамики состоит в том, что десинхронизация различных ансамблей осуществляется очень быстро, так как невзаимодействующие хаотические траектории расходятся с экспоненциальной скоростью.

Заметим, что процедура адаптации векторов $c_i(t)$ существенна. Без нее пиксели со сходными признаками могли бы попасть в разные кластеры. Но стоит заметить, что модель не предусматривает обратного влияния синфазности осцилляторов на процесс адаптации. Таким образом, синфазность лишь выявляет структуру кластеров, формируемых благодаря адаптации.

Пример сегментации изображения с помощью модели Жао приведен на рис. 19.

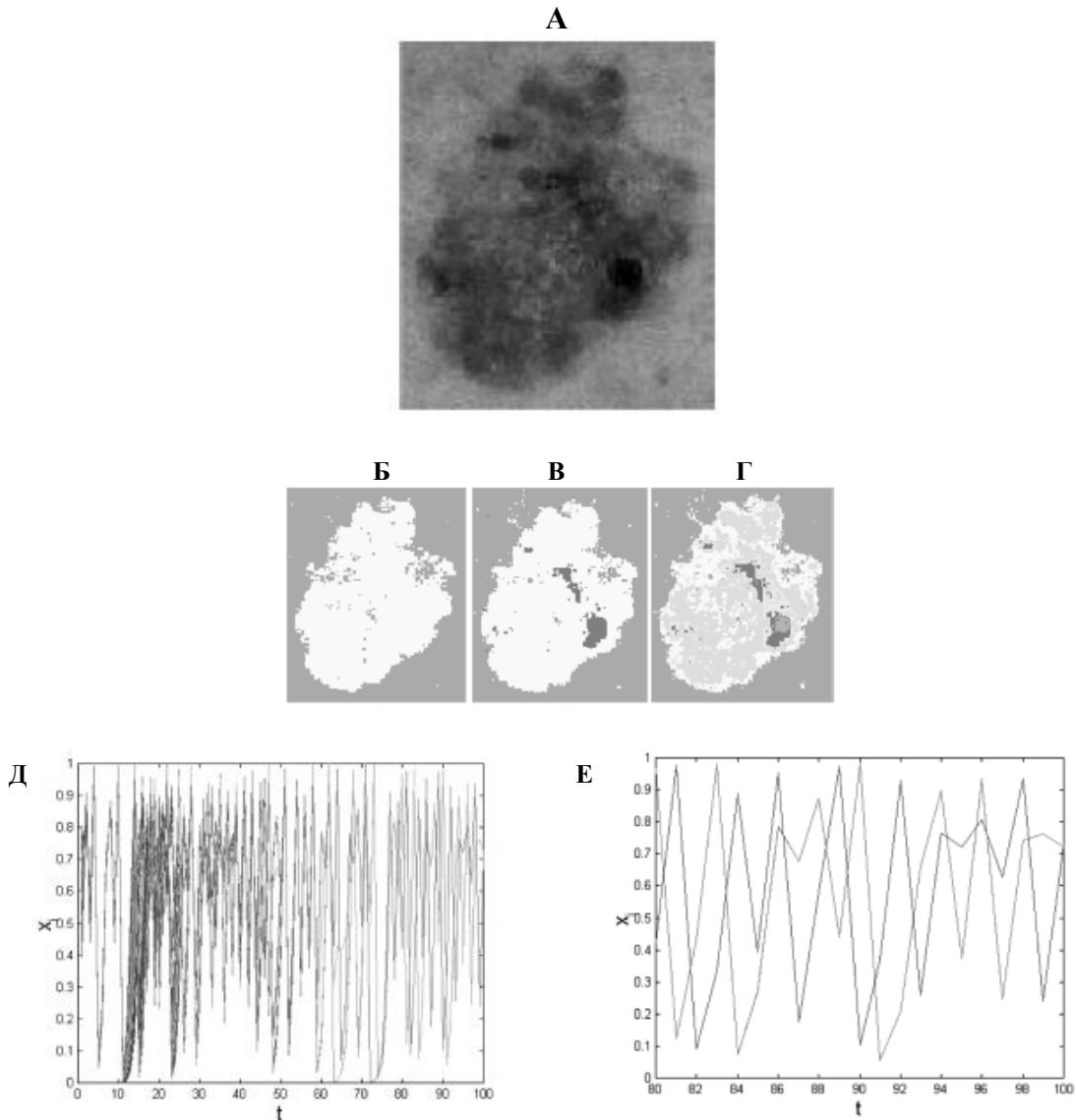


Рис. 19. Пример сегментации изображения моделью Жао [83]. **А.** Исходное изображение 104×123 пикселей. **Б.** Два кластера, $\theta = 0,55$. **В.** Четыре кластера, $\theta = 0,7$. **Г.** Девять кластеров, $\theta = 0,86$. **Д.** Хаотическая динамика осцилляторов, соответствующая кластеризации при $\theta = 0,55$. Несинхронная динамика сменяется синхронной приблизительно с момента времени $t = 50$. **Е.** Фрагмент рисунка Д.

Меняя параметр модели, можно получать различное число сегментов (рис. 19Б - Г). На рис. 19Д - Е представлена динамика модели с разным временным разрешением. Из рис. 19Д видно, что несинхронная динамика сменяется синхронной приблизительно с

момента времени $t = 50$. На рис. 19Е видно, что имеется две несинхронных между собой колебательных активности, соответствующие двум кластерам.

4.3. Модель Буманна-Ланге-Рамахера

В работе Буманна с соавторами [6] предложен еще один вариант сокращения вычислений при сегментации, основанный на том, что элементами, подлежащими сегментации, являются не отдельные пиксели, а квадраты, содержащие сравнительно большое число пикселей. Предположим, что входное изображение разбито на N квадратов одинакового размера. Каждый квадрат i можно характеризовать гистограммой интенсивностей $\mathbf{h}_i = (h_{ij})_{j=1,\dots,M}$ (предполагается, что гистограмма имеет M бинов). Бины гистограммы по определению являются ее признаками. Вероятности признаков определяются как

$$p(j|i) = \frac{h_{ij}}{n}, \quad j = 1, \dots, M,$$

где n - число пикселей в квадрате. Задача сегментации изображения может быть сформулирована как разбиение множества квадратов на K непересекающихся кластеров (сегментов изображения). Предполагается, что величина K известна априорно. Очевидно, разумная процедура кластеризации квадратов должна принимать во внимание как вероятности признаков, так и пространственное расположение квадратов: близлежащие квадраты с большей вероятностью могут принадлежать одному сегменту, чем удаленные квадраты (последнее условие носит название условия локальной гладкости).

Данный подход к сегментации изображений был реализован в модели из импульсных нейронов. Идея модели состоит в том, что веса связей в нейронной сети адаптируются в процессе ее работы таким образом, что после определенного периода обучения сеть оказывается способной классифицировать квадраты в соответствии с некоторой оптимизационной процедурой.

Модель построена из простых импульсных нейронов с постоянным порогом срабатывания. Динамика мембранного потенциала U_i нейрона i задается уравнением

$$\frac{dU_i}{dt} = -\varepsilon U_i + \sum_{k \in \Gamma_i} W_{ki} S_k,$$

где Γ_i - множество пресинаптических нейронов для нейрона i , S_k - текущий выходной сигнал нейрона k , W_{ki} - синаптические веса, ε - параметр затухания мембранного потенциала. В тот момент, когда U_i достигает порога θ , нейрон генерирует импульс длительностью τ , после чего мембранный потенциал немедленно обнуляется.

С каждым квадратом ассоциируется двухслойная сеть с M нейронами во входном слое и K нейронами в выходном (рис. 20). Эта сеть называется блоком. Между слоями блока имеются адаптивные связи, идущие снизу вверх, типа все-на-всех. Кроме того, в выходном слое имеются латеральные тормозные связи типа все-на-всех, реализующие принцип "победитель получает все". Входные нейроны стимулируются сигналами $p(j|i)$ ($j = 1, \dots, M$), так что их импульсная активность пропорциональна указанным вероятностям. Выходные нейроны осуществляют классификацию квадратов. Квадрат i считается принадлежащим k -му сегменту, если k -й выходной нейрон имеет наибольшую активность (частоту разрядов) в выходном слое данного блока.

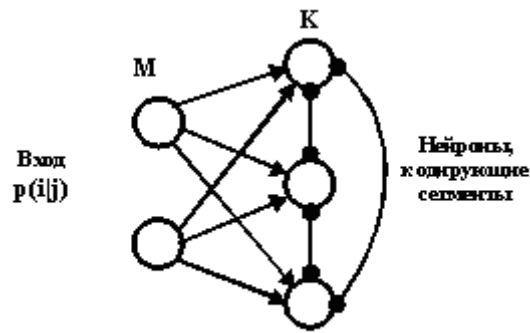


Рис. 20. Архитектура связей в блоке. Стрелками указаны адаптивные возбуждающие связи, а линиями с черными кружками на концах двусторонние тормозные связи. M и K – число нейронов, соответственно, во входном и выходном слоях.

Полная сеть модели строится как плоская сеть из блоков с локальным взаимодействием между блоками. Соседство между блоками соответствует соседству между квадратами. Пусть v_i – выходной нейрон в блоке i , а v_j – выходной нейрон в блоке j . Если блоки i и j являются соседними (в вычислениях использовались ближайшие 4 соседа) и $v_i = v_j$, тогда нейроны v_i и v_j соединены возбуждающими связями, в противном случае связи между нейронами v_i и v_j тормозные. Такая архитектура связей дает некоторое предпочтение отнесению соседних квадратов к одному сегменту, чем повышается гладкость границ сегментируемых областей. Величины связей между блоками фиксированные.

Опишем теперь процедуру обучения (кластеризации). Обозначим $q(j|v)$ вероятность признака j при условии, что квадрат принадлежит сегменту v . Пусть $m_{iv} = 1$, если v – "победивший" сегмент в блоке i , в противном случае $m_{iv} = 0$. Величину

$$\bar{h}_{jv} = \frac{\sum_i m_{iv} p(j|i)}{\sum_i m_{iv}}$$

можно рассматривать как текущую оценку $q(j|v)$. Принимая это во внимание, процедуру адаптации весов связей можно задать следующей формулой:

$$\frac{dW_{jv}^i}{dt} = \alpha (\exp(W_j^{\max} - W_{jv}^i) \bar{h}_{jv} - 1),$$

где α – скорость обучения ($\alpha \approx 0,01$), W_j^{\max} – параметры. Динамика весов связей дает следующее приближенное выражение

$$W_{jv}^i \approx W_j^{\max} + \ln(q(j|v)).$$

Величины W_j^{\max} выбираются таким образом, чтобы гарантировать положительность $W_j^{\max} - W_{jv}^i$. Можно показать, что такой выбор весов связей дает статистически оптимальную схему классификации.

Модель тестировалась при следующих значениях параметров: $N = 64 \times 64$, $M = 12$, $K = 5$ (в одном примере было $K = 2$). Начальные веса и мембранные потенциалы инициализировались случайным образом. Обучение начиналось с небольшим запозданием по отношению к моменту старта функционирования модели, чтобы дать модели время на начальную стабилизацию.

Численные эксперименты проводились с искусственными изображениями, содержащими простые геометрические объекты, искусственные текстуры и реальные изображения. Путем регистрации средней активности нейронов в выходном слое было показано, что 1) нейронная активность имеет осцилляторный характер, 2) нейрон "победившего" класса (сегмента) доминирует над другими нейронами в своем блоке, 3) нейроны, кодирующие один и тот же сегмент, работают синфазно, генерация импульсов нейронами, кодирующими разные сегменты, происходит с фазовым сдвигом. В экспериментах наблюдались волны активации, распространяющиеся по сегментируемым областям.

Результаты сегментации текстур оказались неудовлетворительными, что неудивительно - локальные гистограммы не являются информативными для большинства типов текстур (в данном случае габоровские фильтры были бы более подходящим инструментом).

Пример сегментации реального изображения приведен на рис. 21. На рисунке различные сегменты показаны различными интенсивностями серого цвета.



Рис. 21. Пример сегментации изображения моделью Буманна с соавторами [6]. **А.** Исходное изображение. **Б.** Промежуточные результаты сегментации в момент $t = 5$. **В.** Окончательные результаты сегментации при установлении стационарного состояния сети $t = 50$.

4.4. Обсуждение

Общим недостатком, свойственным всем алгоритмам сегментации, основанным на кластеризации, является то, что эти модели неявно предполагают, что диапазон изменения оптических характеристик в рамках каждого сегментируемого объекта мал по сравнению с различием этих характеристик у разных объектов. Для многих классов изображений это условие не выполняется. В частности, отблески света на поверхности объектов могут существенно повлиять на результат сегментации.

Сравнение моделей Роума-Фригуи и Жао с соавторами показало, что последняя из них работает существенно быстрее, требуя для выполнения кластеризации меньше итераций. Это объясняется с одной стороны тем, что синхронизация, обусловленная непрерывным взаимодействием, наступает быстрее, чем при импульсном взаимодействии. С другой стороны, формирование кластеров ускоряется за счет адаптации входной информации (движения кластеризуемых точек). Вычисления показали, что с увеличением объема обрабатываемых данных число итераций, требуемых для получения компактных кластеров, растет медленно.

Существенный недостаток модели Буманна с соавторами состоит в том, что результат сегментации критически зависит от априорно заданного числа сегментов K . В принципе, можно провести кластеризацию для нескольких значений K , а затем выбрать ту из них, которая окажется более удовлетворительной согласно некоторому дополнительному критерию. Но такой метод проб и ошибок, помимо сомнительной биологической правдоподобности, потребует существенно больших затрат времени на вычисления.

5. СЕГМЕНТАЦИЯ И КОГНИТИВНЫЕ ФУНКЦИИ МОЗГА

Рассматривавшиеся выше методы сегментации практически не включали в себя какой-либо априорной информации относительно того, какие именно объекты должны сегментироваться. Правда, в моделях типа LEGION приоритет при сегментации отдается бóльшим по размеру объектам, а небольшие шумовые объекты могут вообще игнорироваться. Но это весьма частный случай общей ситуации, когда априорно заданная информация о свойствах сегментируемых объектов может существенно повысить качество сегментации. Эта информация может быть задана в достаточно полном объеме в заранее сформированной памяти или путем указания специальных характеристик объектов, таких как яркость, наличие определенных признаков и т.д., которыми должны обладать сегментируемые объекты. Первый из этих случаев соответствует сегментации с памятью, а второй – сегментации с вниманием. Подача на сеть, выполняющую сегментацию, дополнительных сигналов от памяти или внимания способствует созданию неравноправных условий конкуренции для ансамблей осцилляторов, представляющих разные объекты. В результате, победить в конкуренции и сохранить высокую активность способны лишь те ансамбли, которые получают дополнительное подкрепление.

5.1. Сегментация и память

Выше уже указывалось, что сегментация пересекающихся объектов является одной из важных и трудных задач. Рассмотренные методы ее решения базировались на предположении, что пересекающиеся объекты можно разделить благодаря различию их по яркости или цвету. Но что делать в случае, если эти характеристики у объектов идентичны (по крайней мере, идентичны там, где должна проходить граница между объектами)? Например, как разделить частично наложившиеся друг на друга черный треугольник и черный эллипс? Очевидно, такая задача может быть решена только в том случае, если в памяти системы содержится априорная информация об этих геометрических фигурах. Одна из моделей байндинга и сегментации с памятью была предложена в работах Кноблауха и Палма [35 - 37] (краткое изложение модели можно найти также в [55]). В этой модели память реализована в виде отдельного центрального модуля, имеющего прямые и обратные связи с модулем, представляющим первичные области новой коры, где реализуется байндинг. Хотя эта модель работает с контурными изображениями, она рассматривается в данном разделе, поскольку основной ее особенностью является наличие памяти. Нет сомнений, что эта модель может быть адаптирована и для работы с пиксельными изображениями.

Элементами сети являются импульсные нейроны с динамическим порогом срабатывания, сходные с традиционными интегративно-пороговыми нейронами с утечкой [17]. Важным свойством этих нейронов является способность быстро синхронизоваться на временах, соответствующих экспериментальным данным по синхронизации гамма-колебаний. Возбуждающие нейроны в модели обладают относительной рефрактерностью – после генерации импульса порог нейрона мгновенно увеличивается до заданного уровня, а затем экспоненциально падает.

Архитектура модели Кноблауха-Палма представлена на рис. 22. Сеть состоит из трех слоев: сетчатки R , первичной зрительной коры P и центральной области зрительной памяти S . Слой R имеет размер 100×100 нейронов, находящихся во взаимнооднозначном соответствии с пикселями на изображении. Слой P состоит из популяции возбуждающих нейронов P (100×100 нейронов) и двух популяций тормозных нейронов P^S и P^T того же размера (в работе [35] сходные результаты были получены для одной тормозной популяции в P). Аналогично, слой S состоит из трех популяций: возбуждающей популяции S (40×40 нейронов), тормозной популяции для

разделения объектов C^S (40×40 нейронов), тормозной популяции для последовательной смены объектов C^T (20×20 нейронов).

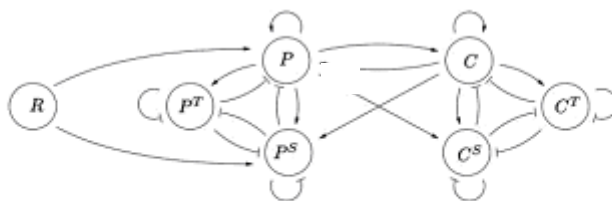


Рис. 22. Архитектура связей в модели Кноблауха-Палма. R – сетчатка, P – первичная зрительная кора, C – память [36].

Поскольку модель ориентирована на работу с контурными изображениями, функционирование слоя P по существу сходно с моделью Ли (см. подраздел 2.1). Отличие данной модели в том, что детекторы направлений расположены не в колонке, как в модели Ли, а на плоскости (детектируемые ориентации ограничены четырьмя направлениями $0^\circ, 45^\circ, 90^\circ, 135^\circ$). Это упрощает архитектуру слоя, но приводит к "рваному" представлению в P непрерывных контуров, активированных в R. Еще одной особенностью модели является введение шума с пространственной и временной корреляцией. Это сделано для ускорения десинхронизации удаленных друг от друга участков контуров и для нарушения неправдоподобно регулярной (с нейрофизиологической точки зрения) колебательной активности.

Слой памяти C представляет собой вариант ассоциативной памяти Виллшоу [79], реализованной на импульсных нейронах. Слой C связан с P прямыми и обратными связями, настраиваемыми с помощью хеббовского обучения и обеспечивающими гетероассоциативность между представлениями стимула в P и C. Связи $C \rightarrow C$ и $C \rightarrow C^S$ являются хеббовскими, остальные связи в C фиксированные. Популяция C^S используется для реализации конкуренции между ансамблями в C, представляющими различные объекты, предотвращая их одновременную активацию. Популяция C^T предназначена для прерывания текущей активности популяции в C, давая возможность активироваться ансамблю, представляющему другой объект. Кроме того, C^T предотвращает неконтролируемое распространение возбуждения в случае, когда силы, разделяющие ансамбли, недостаточно велики.

Рис. 23 показывает, как три тестовых стимула (треугольник, эллипс и квадрат) представляются активностями в слоях R, P и C. Заметим, что в C стимулы представлены бинарными случайными паттернами, которые лишь отдаленно сохраняют форму представления тех же стимулов в R.

Численные расчеты, в которых в качестве входного стимула использовался треугольник, были проведены для сравнения динамики сети в полной модели и в модели, из которой были исключены обратные связи $C \rightarrow P$. В обоих случаях наблюдалась осцилляторная активность в C, но в полной модели частота колебаний в C существенно возросла (с примерно 30 до 50 - 60 герц), а колебания в P стали более регулярными и синфазными с колебаниями в C. Активность всего нейронного ансамбля, представляющего стимул в P, была синфазной. Активность отдельных нейронов в общих чертах отражала активность всего ансамбля, но была намного менее регулярной (например, нейрон в C зачастую срабатывал лишь на каждом втором или третьем колебании суммарной активности популяции). Увеличение суммарной импульсной активности сопровождалось возрастанием частоты появления пачек из двух-трех импульсов на уровне одиночных элементов.

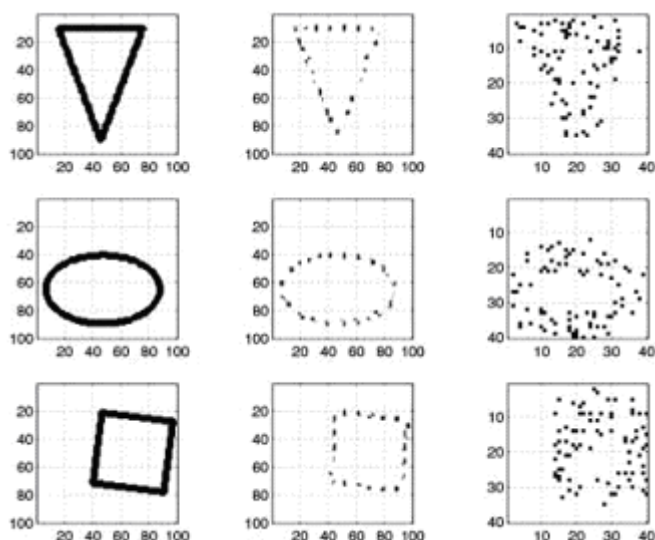
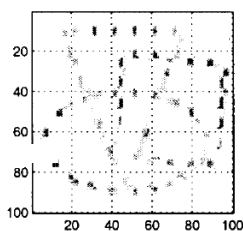


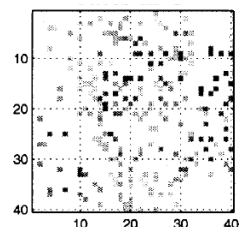
Рис. 23. Представление стимулов в модели Кноблауха-Палма. Слева – представление стимулов на сетчатке R. В центре – представление стимулов в первичной зрительной коре P. Справа – представление стимулов в памяти C [36].

А

Частота импульсации в P



Частота импульсации в С



Б

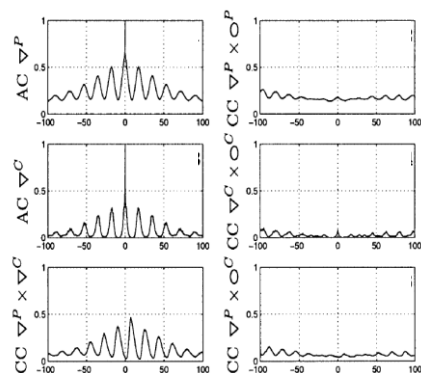


Рис. 24. Функционирование модели Кноблауха-Палма при предъявлении стимула, являющегося суперпозицией трех фигур – треугольника, квадрата и эллипса [36]. **А.** Частоты импульсации нейронов в слоях P и C. **Б.** Автокоррелограммы (AC) и кросс-коррелограммы (CC) суммарной (по каждой фигуре) активности нейронов.

Для тестирования способности модели к сегментации на ее вход был подан стимул, представляющий собой суперпозицию из трех объектов: треугольника, квадрата и эллипса (рис. 24). Предварительно эти объекты предъявлялись поодиночке и были запомнены в ассоциативной памяти.

По сравнению со многими другими моделями сегментации, в которых объекты кодировались синхронными колебаниями со сдвигом фаз между различными ансамблями, в данной модели имела место некоррелированная активность ансамблей, представляющих различные объекты, причем отсутствие корреляции (плоские

коррелограммы) наблюдались в обоих слоях Р и С. Важность этого факта обусловлена тем, что именно такое отсутствие корреляции наблюдалось в нейрофизиологических экспериментах [18, 68].

Благодаря тормозному действию популяции C^S , в каждый момент времени только один ансамбль в С демонстрирует колебательную активность. Для каждого ансамбля период быстрой осцилляторной активности, длившийся несколько сотен миллисекунд (эта длительность зависит от параметра, задающего относительную рефрактерность), сменяется более длительным периодом с низкой нерегулярной активностью. Эти две фазы активности перемежаются на всем протяжении функционирования сети. Аналогичный вид имеет активность в Р. Последнее является результатом действия обратных связей от С к Р.

Исследование коррелограмм активности в Р и С показало, что активность в этих слоях синхронизуется в противофазе при реалистичных временных задержках передачи сигналов (около 4 - 8 мсек). Это противоречит как экспериментальным данным, так и теоретическим представлениями ТВК, согласно которым активность нейронов, представляющих один объект, должна быть синфазной. Авторами предложено несколько вариантов модификации модели для преодоления этой трудности [37]. Один из вариантов был исследован более подробно методами имитационного моделирования. Он заключается во введении бимодальных распределений для синаптических задержек между Р и С. Одна мода представляла собой короткие задержки (около 0,5 мсек), а другая мода порождала задержки, распределенные в широком диапазоне (до 50 мсек). В такой модификации модели была продемонстрирована синфазность колебаний в областях Р и С.

5.2. Сегментация и внимание

Биологические исследования показывают, что обработка зрительной информации у животных и человека не ведется с одинаковой тщательностью по всей зрительной сцене. В большинстве случаев основные усилия сконцентрированы на каких-то специальных областях или объектах зрительного поля. Например, это могут быть области, которые заметно выделяются из своего окружения (по-английски это называется *pop-out*), и объекты, которые только что появились в сцене или представляют в данный момент особую важность. Остальные детали и объекты зрительной сцены более или менее игнорируются. Механизм, позволяющий мозгу осуществлять такую фильтрацию, называется вниманием. Внимание необходимо вследствие ограниченных возможностей зрительной системы, у которой недостаточно ресурсов для детального и быстрого анализа всей зрительной сцены. Важно отметить, что фильтрация зрительной информации может выполняться при неподвижных зрачках, то есть за счет внутренних процессов в мозге. Соответствующее этому внимание называется скрытым (в отличие от открытого внимания, сопровождаемого саккадами - движениями зрачков). В описываемых ниже моделях речь всегда идет о скрытом внимании, поэтому термин "скрытое" будет опускаться.

Модель сегментации с вниманием предложена в работе [5]. Элементами модели являются импульсные нейроны Эхорна [17]. Как уже отмечалось, сети из таких нейронов с локальными связями способны формировать кластеры синфазных осцилляторов. Подача на вход такой сети изображения, представленного интенсивностями пикселей, приведет к сегментации изображения, при которой в каждый кластер будут включены пиксели с близкими значениями интенсивности.

Новой идеей, реализованной в модели, является использование в качестве входной информации при сегментации не только исходного полутонового изображения, но и нескольких изображений, получаемых из исходного в результате его обработки набором фильтров. Назначение фильтра – вычислить в каждом пикселе изображения

значение некоторого признака, ассоциированного с данным фильтром. Полученное в результате фильтрации изображение называется картой признака. В реальных вычислениях использовались три типа фильтров: вейвлеты, морфологические фильтры, характеризующие локальную геометрию изображения, и разностно-гауссовские фильтры (difference-of-Gaussians - DoG фильтры), поскольку эти фильтры уже зарекомендовали себя при анализе различных классов изображений [29, 57, 60]. Использование карт признаков позволяет найти на изображении наиболее вероятную область расположения целевых объектов. Благодаря ограничению области, где должна проводиться сегментация, качество сегментации с помощью временной корреляции удастся повысить. На выходе модели сегментируемые объекты представлены нейронными ансамблями с более высоким уровнем активности.

Модель состоит из взаимодействующих модулей: центрального модуля и набора модулей признаков (рис. 25). На вход центрального модуля подаются интенсивности пикселей. На вход модуля признака подаются сигналы от карты этого признака. Нейроны внутри модуля организованы в виде двумерного слоя с локальными связями (в вычислениях связи имели радиус 3).



Рис. 25. Схема модели сегментации с вниманием, реализованная в виде импульсной нейронной сети [5].

Сегментация на основе временной корреляции выполняется центральным модулем. Роль, исполняемая модулями признаков, сходна с той, которую играет карта заметности в модели внимания Итти и Коха [27 - 28]. Модули признаков подают на центральный модуль дополнительные возбуждающие модулирующие сигналы, цель которых сфокусировать внимание на сегментах изображения, представляющих интерес. В результате активность нейронов, представляющих целевые объекты, оказывается существенно выше активности нейронов, представляющих нецелевые объекты и фон.

В качестве тестовых примеров для проверки качества работы модели использовались маммограммы и инфракрасные изображения ракетных установок. В случае ракетных установок модель позволяла отфильтровать 94% объектов, не являющихся целевыми, при нулевых ошибках распознавания целевых объектов. В случае маммограмм модель позволяла отфильтровать 46% нецелевых объектов при 7% ошибок распознавания целевых объектов. Такого качества сегментации невозможно было бы добиться на основе какого-либо одного метода фильтрации по заданному признаку или их простой логической комбинацией.

Работа [5] была ориентирована главным образом на техническое применение идей ТВК; вероятно, поэтому в ней отсутствует корреляционный анализ режимов синхронизации, возникающих при функционировании модели. В результате остается неясным, имеется ли десинхронизация активности нейронных ансамблей, кодирующих различные объекты на изображении, и если да, то какими механизмами эта десинхронизация обеспечивается.

В работе [40] используются сходные идеи, но при этом проблема синхронизации/десинхронизации исследовалась явно. Интенсивности исходного изображения представлялись во входном слое модели амплитудами осцилляторов ФитцХью-Нагумо. Для реализации десинхронизации был введен специальный "возмущающий" слой, десинхронизирующие сигналы с которого подавались на входной слой. Это давало требуемый эффект десинхронизации, но, к сожалению (как было выяснено кросс-корреляционным анализом), активность пары нейронов, представляющих разные объекты, оказалась сильно коррелированной. Это означает, что данные нейроны генерируют последовательности импульсов, отличающиеся всего лишь сдвигом фаз. Такая активность противоречит имеющимся экспериментальным данным [18, 68].

Модели сегментации [5, 40] основаны на автоматической фокусировке внимания по принципу pop-out. Другой возможный подход к реализации внимания основан на неравноправной конкуренции (biased competition) между нейронными ансамблями, когда вводится дополнительный сигнал, повышающий активность какого-либо ансамбля и тем самым обеспечивающий его "победу" в конкуренции по принципу "победитель получает все". Считается, что такой дополнительный сигнал вырабатывается системой внимания в высоких областях новой коры и затем проектируется в первичные зоны зрительной коры, формируя подходящий фокус внимания.

В уже рассматривавшихся выше работах [35 - 37, 55] (см. раздел 5.1) авторы использовали неравноправную конкуренцию для проверки способности своей модели воспроизводить экспериментальные данные по регистрации активности нейронов, стимулируемых комбинациями объектов [50, 58]. Сначала в памяти модели были запомнены два объекта: квадрат и эллипс. Затем на вход модели подавались каждый из этих стимулов по отдельности или их комбинация (сложный стимул). Регистрировалась (в виде постстимульной гистограммы с временным шагом 500 мсек) средняя активность нейрона, чувствительного к стимулу "квадрат". В согласии с экспериментальными данными при моделировании были получены следующие результаты:

- 1) при отсутствии внимания ответ нейрона на сложный стимул был меньшим, чем на "свой" стимул (квадрат), и большим, чем на "чужой" стимул (эллипс);
- 2) при наличии внимания ответ нейрона на сложный стимул был таким, как если бы предъявлялся только один объект – тот, на котором было сфокусировано внимание.

Указанные эффекты были более выражены в верхнем слое памяти (слой С), чем в нижнем слое первичной зрительной коры (слой Р).

Механизм, благодаря которому модель правильно воспроизводит экспериментальные данные, связан с участием нейрона, чувствительного к определенному стимулу, в двух режимах байндинга - с высокой и низкой активностью. При предъявлении "своего" стимула нейрон все время находится в состоянии высокой активности. При предъявлении сложного стимула и при отсутствии внимания нейрон примерно половину времени проводит в состоянии высокой активности, а в остальное время находится в состоянии низкой активности. Если же внимание на "своем" стимуле присутствует, нейрон снова возвращается в режим высокой активности в течение всего времени стимуляции.

В разделе 1 были описаны экспериментальные данные, поддерживающие гипотезу о наличии центрального элемента системы внимания, локализованного в префронтальной коре и являющегося источником синхронизации и байндинга в зрительной коре [24]. Идея центрального элемента системы внимания была положена в основу осцилляторной модели внимания и сегментации, реализованной с использованием пейсмекерных импульсных нейронов типа Ходжкина-Хаксли в качестве элементов модели [11, 95].

Достоинство данной модели в том, что она детально отражает функционирование нейронов. Это позволяет исследовать режимы синхронизации импульсной активности, ассоциированные с вниманием, воспроизводить подавление импульсной активности, соответствующей стимулам вне фокуса внимания, выбирать в фокус внимания наиболее заметные стимулы и воспроизводить частотные характеристики нейронной активности, наблюдаемые в экспериментах по вниманию.

Архитектура модели представлена на рис. 26. Элементы первого слоя называются периферическими нейронами (ПН). Они представляют собой детекторы признаков в первичных зонах зрительной коры, активируемые внешними стимулами. Предполагается, что внешний вход на ПН достаточно велик, чтобы перевести нейрон в пейсмекерный режим с некоторой частотой генерации импульсов. В случае зрительных стимулов входные сигналы формируются изображением, расположенным на плоской решетке. ПН расположены на решетке с такими же размерами, при этом каждый ПН получает сигнал от пикселя, имеющего те же координаты на решетке, что и ПН. Вообще говоря, ПН должны быть связаны возбуждающими латеральными связями, но для упрощения анализа системы эти связи элиминированы, так что все взаимодействие между ПН идет через центральный элемент.

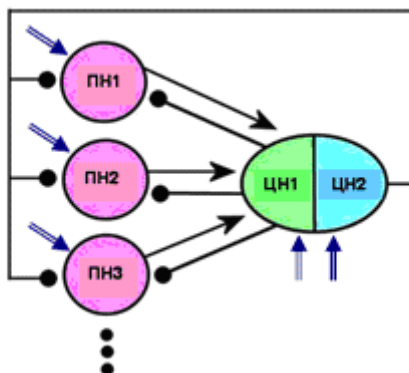


Рис. 26. Архитектура связей в модели внимания с центральным элементом. ПН1, ПН2, ПН3,... – периферические нейроны, кодирующие признаки внешних стимулов; ЦН1, ЦН2 – центральные нейроны, управляющие фокусом внимания. Стрелками показаны возбуждающие связи, линии с кружками на концах показывают тормозные связи, двойные синие стрелки показывают внешние входные сигналы.

Центральный элемент состоит из двух нейронов, ЦН1 и ЦН2, и представляет собой предельно упрощенный вариант центрального элемента системы внимания. ЦН1 используется для синхронизации некоторого ансамбля ПН. Зрительный объект, представленный ансамблем ПН, работающим синфазно с ЦН1, считается включенным в фокус внимания (это соответствует экспериментальным данным [24], см. раздел 1). ЦН2 управляет последовательной синхронизацией ЦН1 с разными ансамблями ПН, что интерпретируется как последовательный выбор в фокус внимания разных объектов. Оба нейрона ЦН1 и ЦН2 посылают на все ПН тормозные сигналы. Связи от ПН приходят только на ЦН1 и являются возбуждающими. Сила синаптической передачи предполагается фиксированной и одинаковой для всех таких связей. Связи между ЦН1

и ЦН2 отсутствуют. Таким образом, ЦН2 не получает никаких сигналов от других нейронов системы. На вход ЦН2 приходит только постоянный внешний сигнал, переводящий ЦН2 в осцилляторный режим.

В то время как сила торможения ПН со стороны ЦН1 фиксирована, тормозные сигналы, поступающие на ПН от ЦН2, меняются со временем. Последнее обусловлено кратковременной синаптической пластичностью связей от ЦН2 на ПН. Если оба нейрона, ЦН2 и ПН, одновременно активны достаточно продолжительное время, сила тормозной синаптической связи между ними возрастает. Если же хотя бы один из этих нейронов имеет низкий уровень активности, сила синаптической связи падает. Данный способ адаптации тормозной связи от ЦН2 к ПН используется для моделирования экспериментального факта, согласно которому при последовательном просмотре объектов вероятность повторного выбора уже просмотренного объекта падает [34].

Параметры взаимодействия между ПН и ЦН1 выбраны таким образом, что ЦН1 способен генерировать импульсы даже при отсутствии внешнего входа. Наличие возбуждающего внешнего входа приводит к увеличению частоты импульсации ЦН1.

С помощью связей от ЦН1 к слою ПН происходит торможение импульсной активности тех ПН, которые в данный момент не включены в фокус внимания. Сила связи от ЦН1 к ПН предполагается постоянной и одинаковой для всех ПН.

Синаптическая связь от ЦН2 к ПН имеет низкое по абсолютной величине значение, если данный ПН не включен в фокус внимания, и скачком увеличивается до высокого по абсолютной величине значения через некоторое время после того, как ПН оказывается включенным в фокус внимания. Это высокое значение удерживается определенное время, а затем сила связи скачком снова возвращается к низкому значению. Такое правило модификации ограничивает время, в течение которого объект может находиться в фокусе внимания. Кроме того, оно не позволяет объекту, уже включавшемуся ранее в фокус внимания, снова оказаться в фокусе внимания прежде, чем через фокус внимания будут пропущены другие объекты. В результате система внимания последовательно переключается с одного объекта на другой, не застревая на наиболее заметном объекте.

Согласно постулатам модели считается, что объект, представленный нейронами $ПН_{i1}, \dots, ПН_{ik}$, включен в фокус внимания, если активность этих нейронов когерентна активности нейрона ЦН1, а импульсная активность остальных нейронов заторможена. Такой режим функционирования модели называется режимом частичной синхронизации.

На рис. 27 показана динамика мембранных потенциалов нейронов в режиме частичной синхронизации. в двух группах, различающихся собственными частотами осцилляторов.

Группа нейронов ПН1, ..., ПН5 генерирует импульсы синфазно с центральным нейроном ЦН1. Активность нейронов в группе ПН6, ..., ПН10 подавлена, так что эти нейроны не могут генерировать импульсы.

Компьютерное моделирование показало, что для наступления частичной синхронизации с одной из групп осцилляторов основным фактором является наличие более высоких собственных частот осцилляторов в этой группе по сравнению с собственными частотами осцилляторов в другой группе. С точки зрения нейробиологии этот факт важен потому, что более высокая заметность стимула приводит к более высокой активности нейронов, представляющих этот стимул в коре мозга, и к большей вероятности выбора этого стимула в фокус внимания.

Пример, показывающий способность модели к сегментации с последовательным выбором объектов представлен на рис. 28. Изображение, содержит три объекта - синюю салфетку, оранжевый апельсин и желтую грушу - на фоне белой скатерти. Различные цвета пикселей на изображении были закодированы различными собственными

частотами ПН. Для визуализации области текущей фокусировки внимания, использован бежевый цвет, т.е. в момент, когда некоторый ПН генерирует импульс, соответствующий пиксель на рис. 28 окрашивается в бежевый цвет.

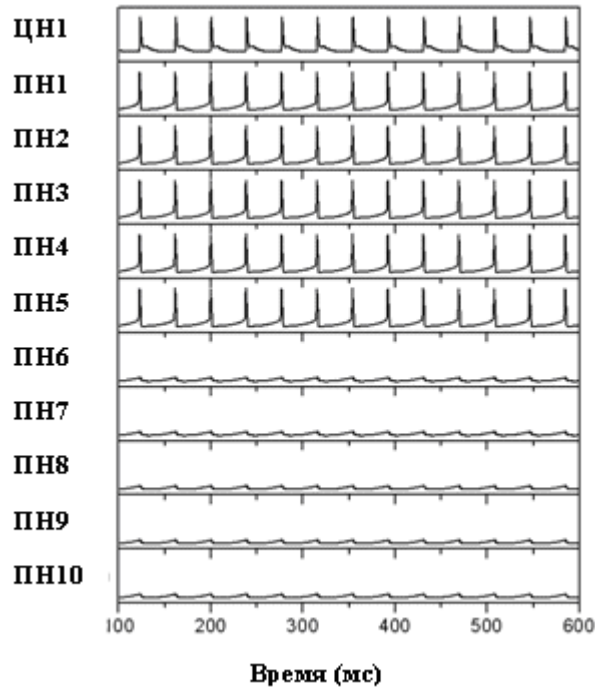


Рис. 27. Динамика мембранных потенциалов в двух группах нейронов в режиме частичной синхронизации [11].

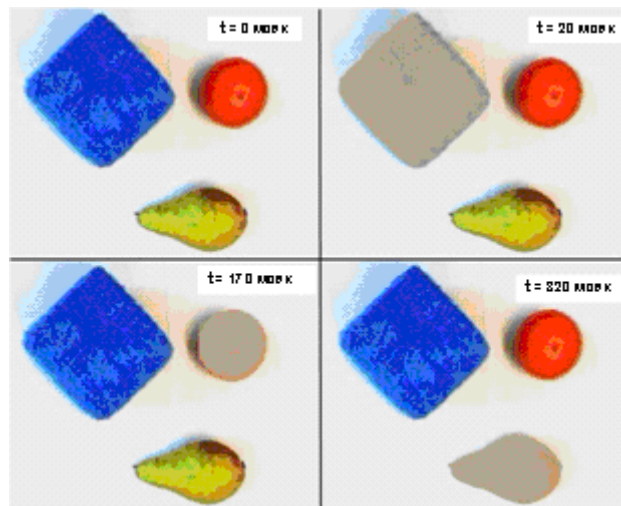


Рис. 28. Последовательный выбор в фокус внимания объектов на реальном изображении. Слева сверху – исходное изображение, справа сверху – выбор салфетки, слева внизу – выбор апельсина, справа внизу – выбор груши [11].

Временной интервал 0–20 мсек является переходным периодом, когда происходит установление фокуса внимания. В момент времени $t = 20$ мсек происходит выбор салфетки в фокус внимания, т.е. в режим частичной синхронизации с ЦН1 входят нейроны, расположенные в области, занимаемой салфеткой. Активность остальных ПН подавляется. Этот фокус внимания сохраняется до момента $t = 170$ мсек, когда происходит выбор апельсина в фокус внимания. Выбор груши в фокус внимания

происходит в момент $t = 320$ мсек. После этого фокус внимания возвращается к салфетке, и т.д.

В любой момент времени одна из групп ПН, представляющих какой-то из имеющихся на изображении объектов, имеет более высокие частоты импульсации, и это обеспечивает включение этой группы в частичную синхронизацию с ЦН1 и, соответственно, фокусировку внимания на данном объекте. Перемещение фокуса внимания обеспечивается нейроном ЦН2. В период тормозного воздействия ЦН2 на группу ПН импульсная активность этой группы подавляется. Группа возвращается к импульсной активности и получает возможность работать в режиме частичной синхронизации после того, как тормозное воздействие на нее со стороны ЦН2 прекращается.

Недостаток описанной модели состоит в том, что она работоспособна при контрастных цветах объектов и фона. При наличии на изображении изолированных объектов, окрашенных в похожие цвета, все эти объекты будут одновременно выбираться в фокус внимания, поскольку все осцилляторы, соответствующие этим объектам будут иметь сходные собственные частоты.

Для преодоления этого недостатка была разработана модель со сходной архитектурой, но в которой сегментация выполняется в два этапа. На первом этапе проводилась "грубая" сегментация стандартными средствами искусственного зрения. На втором этапе эта сегментация уточнялась с использованием осцилляторной нейронной сети с центральным элементом [4].

В качестве элементов сети использовались обобщенные фазовые осцилляторы, состояние которых описывается двумя переменными, фазой и амплитудой колебаний. Осцилляторы взаимодействуют по принципу фазовой автоподстройки частоты. При этом амплитуды меняются в зависимости от соотношения фаз центрального осциллятора и периферического осциллятора. Амплитуда возрастает, если эти осцилляторы синфазны и падает по мере рассогласования фаз. Это позволяет моделировать увеличение активности нейронных ансамблей, представляющих объект в фокусе внимания, и подавление активности вне фокуса внимания. Как и в модели [11] фокус внимания представлен ансамблем осцилляторов, работающих синфазно с центральным осциллятором.

Архитектурно модель представляет собой двухслойную сеть периферических осцилляторов с локальными связями внутри слоев и между слоями и с одним центральным осциллятором, управляющим динамикой периферических осцилляторов. Синхронизирующие и десинхронизирующие связи между центральным и периферическими осцилляторами организованы таким образом, что ансамбли осцилляторов, представляющие различные зрительные объекты, конкурируют за синхронизацию с центральным осциллятором. Ансамбль, выигравший конкуренцию, представляет объект, на котором в текущий момент времени сфокусировано внимание.

Предварительная сегментация изображения делается с помощью стандартных фильтров для выделения границ. Эти фильтры дают приблизительное положение границ и некоторое количество шумовых и ложных объектов. Параметры фильтров выбраны так, чтобы в результате фильтрации получить замкнутые границы сегментируемых объектов. Эти границы используются в первом слое модели для разделения процессов синхронизации внутри и вне объекта, который требуется включить в фокус внимания. Это реализуется путем блокировки (остановки) активности осцилляторов, лежащих на границах.

Второй слой модели служит для элиминации границ, шумов и ложных объектов. Осцилляторы второго слоя конкурируют между собой за синхронизацию с осцилляторами первого слоя. В результате во втором слое формируется два ансамбля осцилляторов. Один из них работает синфазно с центральным осциллятором, а другой

сдвинут по фазе относительно центрального осциллятора. Величина фазового сдвига может быть разной для разных осцилляторов и меняющейся со временем.

Пример сегментации объекта в зрительной сцене и включения его в фокус внимания приведен на рис. 29 - 30.

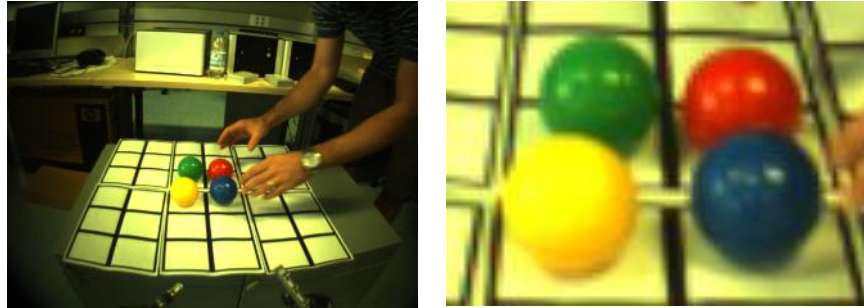


Рис. 29. Исходное изображение (слева) и фрагмент изображения, который подвергался сегментации (справа).

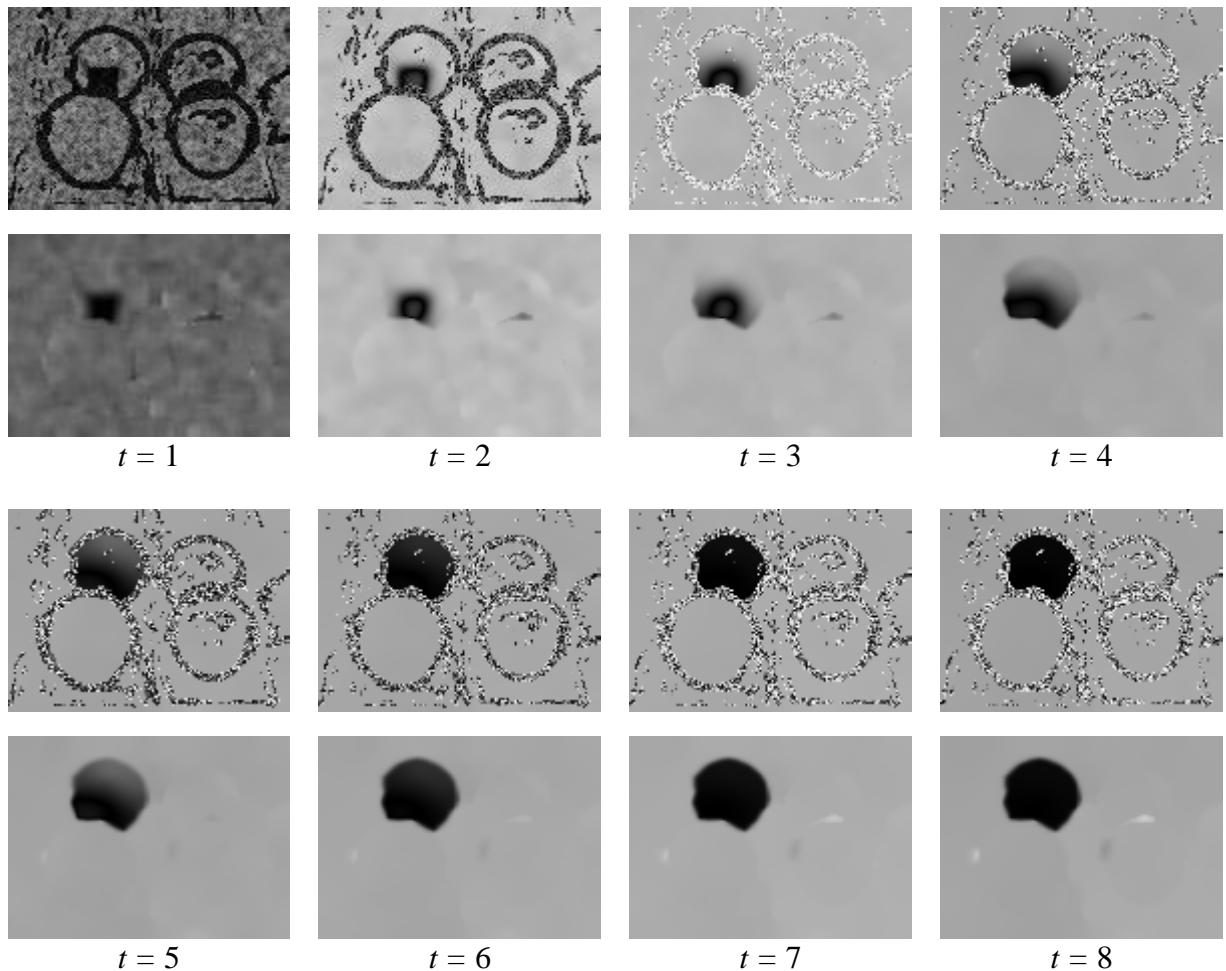


Рис. 30. Сегментация зеленого шара. Каждый пиксель на изображении показывает разность фаз между периферическим и центральным осцилляторами. Разности фаз в диапазоне $(0, 2\pi)$ для визуализации шкалированы в диапазон $(0, 256)$. Нулевая разность фаз соответствует черному цвету. Время дано в условных единицах. Для каждого момента времени верхнее изображение соответствует состоянию первого слоя модели, а нижнее – состоянию второго слоя.

На рис. 29 показана зрительная сцена, анализ которой должен привести к возможности выбора шара определенного цвета рукой робота. Область локализации

шаров выделяется с помощью модели, описанной в [11, 95]. Эта область изображена в правой части рис. 30. Далее этот фрагмент подвергался обработке с помощью данной модели.

Динамика процесса сегментации показана для случая выбора зеленого шара на рис. 30. Для указания, какой именно шар должен выбираться в фокус внимания, на изображении "подсвечивалась" небольшая область заданного цвета (в данном случае зеленого. Эта "подсветка" осуществлялась путем увеличения взаимодействия между осцилляторами в данной области и центральным осциллятором. В результате возникла синхронизация между центральным осциллятором и осцилляторами данной области.

Это показано на рис. 30 в момент времени $t = 1$. Область "подсветки" имеет вид небольшой квадратной области, имеющей черный цвет и расположенной внутри зеленого шара.

В более поздние моменты времени синхронизация постепенно распространяется на окружающие осцилляторы. В первом слое это распространение ограничивается вычисленной предварительно границей объекта. Во втором слое все осцилляторы активны и включены в процесс фазовой синхронизации, в том числе и осцилляторы, лежащие на границе. В результате постепенно формируется синхронный (синфазный) ансамбль осцилляторов, представляющий зеленый шар. Остальные осцилляторы работают со сдвигом фазы по отношению к центральному осциллятору, что приводит к подавлению их активности.

5.3. Обсуждение

Привлечение для сегментации изображений когнитивных функций мозга является перспективным направлением исследований. Экспериментальные данные о том, в какой мере и на каком этапе память и внимание участвуют в реализации сегментации, остаются весьма фрагментарными. По-видимому, простые задачи сегментации могут быть решены автономно на уровне первичных зон зрительной коры, но в более сложных случаях когнитивные функции (память, распознавание, внимание) оказываются необходимыми для успешного решения задачи. При этом когнитивные функции могут участвовать в выполнении сегментации как синхронно с сегментацией, так и итеративно. В последнем случае промежуточные результаты сегментации анализируются на уровне когнитивных функций с привлечением верхних зон коры, а затем полученная таким образом информация используется в первичных зонах зрительной коры для уточнения сегментации.

Выполненные работы показали, что дополнительная информация об объектах, содержащихся в памяти или обладающих специальными свойствами, может существенно повысить качество сегментации, а также дать объяснение ряду нейрофизиологических данных. Однако в существующих моделях пока что не удалось сочетать хорошее соответствие экспериментальным данным с сегментацией естественных изображений. Для естественных изображений сегментация налагающихся объектов с одинаковыми оптическими характеристиками по-прежнему остается нерешенной задачей. Модель сегментации Кноблауха-Палма в настоящем виде не обладает инвариантностью по отношению к положению запомненных объектов. Поэтому она не способна сочетать сегментацию объекта с его поиском на изображении. Достижение инвариантности при взаимодействии сегментации и памяти - важная задача для будущих исследований.

Достоинство моделей [4, 11] состоит в том, что они ориентированы на работу с реальными цветными изображениями и достаточно полно отражают имеющиеся экспериментальные данные о принципах работы зрительной системы и системы внимания. Однако требование замкнутости контуров, получаемых при предварительной фильтрации изображения, является весьма ограничительным для

многих реальных изображений. Требуется разработка более совершенной модели, в которой этот недостаток был бы устранен.

6. ЗАКЛЮЧЕНИЕ

Обзор дает довольно полное представление о моделях сегментации с помощью динамических и осцилляторных нейронных сетей, разработанных в последние годы. Рассмотренные модели показывают, что принципы ТВК являются эффективным средством решения задач сегментации изображений. В существующих моделях удалось воспроизвести ряд экспериментальных данных, а лучшие из разработок продемонстрировали качество сегментации, сравнимое или даже несколько превосходящее системы искусственного интеллекта.

Следует, однако, отметить, что круг воспроизводимых в моделях нейробиологических и психологических данных не слишком широк, а с точки зрения технических приложений существующие системы еще весьма далеки от возможностей сегментации, демонстрируемых человеческим мозгом. Модели используют весьма широкий спектр принципов, что можно считать как достоинством, так и недостатком. Достоинство можно видеть в том, что у экспериментаторов будет больший выбор при планировании и объяснении результатов экспериментов. Недостаток же состоит в том факте, что на данном этапе исследований трудно сделать окончательные выводы относительно того, какие из этих принципов действительно реализуются в биологических организмах. Настораживает и тот факт, что стремление получить высокое качество сегментации изображений зачастую сопряжено с игнорированием известных экспериментальных данных.

Очевидно, что существующие модели далеко не используют те возможности, которые предоставляются имеющимися знаниями о функционировании мозга. Попробуем перечислить несколько направлений, в которых могло бы пойти усовершенствование моделей сегментации.

Практически все модели реализуют байндинг каждого объекта на какой-то одной частоте (например, из гамма-диапазона). Однако иерархическая организация коры мозга и множественность частот, одновременно используемых при обработке зрительной информации, позволяют предположить, что в реализации байндинга может участвовать сразу несколько частот из разных частотных диапазонов. Признаки низкого уровня, возможно, связываются синхронизацией колебаний в диапазоне гамма-ритма, а окончательное связывание признаков в образ сложного объекта происходит при синхронизации на более низкой частоте (например, на частоте тета-ритма).

Хорошо известно, что в мозге приматов зрительная информация после первичных зон зрительной коры разбивается на два потока, один из которых идет в височную кору и отвечает за распознавание и запоминание объектов, а другой идет в теменную кору и отвечает за расположение объектов в пространстве. Эта особенность обработки информации в мозге не нашла отражения в моделях. По-видимому, с этим связаны трудности сопряжения сегментации с другими когнитивными функциями. Взаимодействие различных когнитивных процессов – сегментации, памяти, внимания, распознавания, детекции новизны, в которых задействованы многие структуры мозга, остается за пределами сегодняшних возможностей моделей. Только при создании моделей с широким спектром когнитивных функций можно надеяться получить результаты сегментации, заметно превосходящие системы искусственного интеллекта.

Еще одним резервом улучшения моделей сегментации является использование бинокулярного зрения. Принципы ТВК могут оказаться весьма полезными при интеграции признаков, полученных в разделенных оптических каналах.

С точки зрения решаемых задач, возможно, наиболее перспективными являются задачи по сегментации движущихся изображений. В этой области накоплен большой

экспериментальный материал. Практическая важность таких работ также не вызывает сомнений. Небольшое число имеющихся здесь моделей не вышли на уровень работы с достаточно сложными изображениями и не включены в настоящий обзор. Тем не менее, можно ожидать, что задача слежения за объектами в зрительной сцене, включающей множественные целевые объекты, объекты-дистракторы (отвлекающие внимание) и скрывающие экраны, станет важным стимулом для дальнейшего развития и проверки работоспособности моделей сегментации.

СПИСОК ЛИТЕРАТУРЫ

1. Baddeley A. Exploring the central executive. *Quarterly Journal of Experimental Psychology*. 1996. V. **49A**. P. 5-28.
2. Baddeley A. Fractionating the central executive. In: *Principles of Frontal Lobe Function*, Eds. Stuss D. and Knight R.T. New York: Oxford University Press, 2002. P. 246–260.
3. Borisyuk R., Kazanovich Y. Oscillatory model of attention-guided object selection and novelty detection. *Neural Networks*. 2004. V. **17**. P. 899-915.
4. 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**. P. 707-719.
5. Broussard R.P., Rogers S.K., Oxley M.E., Tarr G.L. Physiologically motivated image fusion for object detection using a pulse coupled neural network. *IEEE Trans. Neural Networks*. 1999. V. **10**. P. 554-563.
6. Buhmann J.M., Lange T., Ramacher U. Image segmentation by networks of spiking neurons. *Neural Computation*. 2005. V. **17**. P. 1010-1031.
7. Caelli T., Bischof W. Machine learning and image interpretation - Plenum Press. New York, 1997.
8. Carson C., Belongie S., Greenspan H., Malik J. Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 2002. V. **24**. P. 1026-1038.
9. Chen K., Wang D.L. A dynamically coupled neural oscillator networks for image segmentation. *Neural Network*. 2002. V. **15**. P. 423-439.
10. Chen K., Wang D.L., Liu X. Weight adaptation and oscillatory correlation for image segmentation. *IEEE Trans. Neural Networks*. 2000. V. **11**. P. 1106-1123.
11. Chik D., Borisyuk R., Kazanovich Y. Selective attention model with spiking elements. *Neural Networks*. 2009. V. **22**. P. 890-900.
12. Choe Y., Miikkulainen R. Self-organization and segmentation in a laterally connected orientation map of spiking neurons. *Neurocomputing*. 1998. V. **21**. P. 139-57.
13. Choe Y., Miikkulainen R. Contour integration and segmentation with self-organized lateral connections. *Biol. Cybern.* 2004. V. **90**. P. 75-88.
14. Cowan N. Evolving conceptions of memory storage, selective attention and their mutual constraints within the human information processing system. *Psychological Bulletin*. 1988. V. **104**. P. 163-191.
15. Doesburg S.M., Roggeveen A.B., Kitajo K., Ward L.M. Large-scale gamma-band phase synchronization and selective attention. *Cerebral Cortex*. 2008. V. **18**. P. 386-396.
16. Eckhorn R., Bauer R., Jordon W., Brosch M., Kruse W., Munk M., Reitboeck H.J. Coherent oscillations: a mechanism of feature linking in the visual cortex. *Biol. Cybern.* 1988. V. **60**. P. 121-130.
17. Eckhorn R., Reitboeck H.J., Amdt M., Dicke P. Feature linking via synchronization among distributed assemblies: Simulation of results for cat visual cortex. *Neural Computation*. 1990. V. **2**. P. 293-307.

18. Engel A., König P., Singer W. Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl. Acad. Sci. (USA)*. 1991. V. **88**. P. 9136-9140.
19. Fries P., Reynolds J., Rorie A., Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*. 2001. V. **291**. P. 1560-1563.
20. Fries P., Schroeder J-H., Roelfsema P.R., Singer W., Engel A.K. Oscillatory neural synchronization in primary visual cortex as a correlate of stimulus selection. *J. Neurosci*. 2002. V. **22**. P. 3739-3754.
21. Gray C.M. Synchronous oscillations in neural systems: Mechanisms and functions. *J. Comput. Neurosci*. 1994. V. **1**. P. 11-38.
22. Gray C.M. The temporal correlation hypothesis of visual feature integration is still alive and well. *Neuron*. 1999. V. **24**. P. 31-47.
23. Gray C.M., König P., Engel A.K., Singer W. Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature*. 1989. V. **388**. P. 334-337.
24. Gregoriou G.G., Gotts S.J., Zhou H, Desimone R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*. 2009. V. **324**. P. 1207-1210.
25. Haralick R.M., Shapiro L.G. Image segmentation techniques. *Computer Vision, Graphics and Image Processing*. 1985. V. **29**. P. 100-132.
26. Herrman C.S., Munk M.H.J., Engel A.K. Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences*. 2004. V. **8**. P. 347-355.
27. Itti L., Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*. 2000. V. **40**. P. 1489-506.
28. Itti L., Koch C. Computational modeling of visual attention. *Nature Reviews Neuroscience*. 2001. V. **2**. P. 194-203.
29. Jin Y., Angelini E., Laine A. Wavelets in Medical Image Processing: Denoising, Segmentation, and Registration. In *Handbook of Biomedical Image Analysis. Vol. 1, Segmentation models, part A (Topics in Biomedical Engineering. International Book Series)*. Eds. Suri J.S., Wilson D.L., Laxminarayan S. Springer, 2005. P. 305-358.
30. Jutras M.J., Gamma-band synchronization in the macaque hippocampus and memory formation. *J. Neurosci*. 2009. V. **29**. P. 12521-12531.
31. Kaneko K. Clustering, coding switching, hierarchical ordering, and coding in a network of chaotic elements. *Physica D*. 1990. V. **41**. P. 137-72.
32. Kapadia M.K, Ito M., Gilbert C.D., Westheimer G. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*. 1995. V. **15**. P. 843-56.
33. Kazanovich Y., Borisyuk R. Object selection by an oscillatory neural network. *BioSystems*. 2002. V. **67**. P. 103-111.
34. Klein R. Inhibitory tagging system facilitates visual search. *Nature*. 1988. V. **334**. P. 430-431.
35. Knoblauch A., Palm G. Pattern separation and synchronization in spiking associative memories and visual areas. *Neural Networks*. 2001. V. **14**. P. 763-780.
36. Knoblauch A., Palm G. Scene segmentation by spike synchronization in reciprocally connected visual areas. I. Local effects of cortical feedback. *Biol. Cybern*. 2002. V. **87**. P. 151-167.
37. Knoblauch A., Palm G. Scene segmentation by spike synchronization in reciprocally connected visual areas. II. Global assemblies and synchronization on larger space and time scales. *Biol. Cybern*. 2002. V. **87**. P. 168-184.
38. Kryukov V.I. An attention model based on the principle of dominantia. In: *Neurocomputers and attention I: neurobiology, synchronisation and chaos*. Eds. Holden A.V., Kryukov V.I. Manchester: Univ. Press, 1991. P. 319-351.

39. Kuzmina M., Manykin E., Surina I. Oscillatory networks with self-organized dynamical connections for synchronization-based image segmentation. *BioSystems*. 2004. V. **76**. P. 43-53.
40. Labbi A., Milanese R., Bosch H. Visual object segmentation using FitzHugh-Nagumo oscillators. *Nonlinear Analysis. Theory, Methods & Applications*. 2001. V. **47**. P. 5827-5838.
41. Li Z. A neural model of contour integration in the primary visual cortex. *Neural Computation*. 1998. V. **10**. P. 903-940.
42. Li Z. Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network*. 1999. V. **10**. P. 187-212.
43. Li Z. Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*. 2000. V. **13**. P. 25-50.
44. Li Z. Computational design and nonlinear dynamics of a recurrent network model of the primary visual cortex. *Neural Computation*. 2001. V. **13**. P. 1749-1780.
45. Li J., Gray R.M. Image Segmentation and Compression Using Hidden Markov Models. In: *The International Series in Engineering and Computer Science*. Kluwer Academic Publishers, 2000.
46. Lindblad T., Kinser J.M. *Image processing using pulse-coupled neural networks*, 2nd ed. Springer-Verlag, 2005.
47. Malsburg C. von der. The correlation theory of brain function. Internal report 81-2, Max-Planck Institute for Biophysical Chemistry, 1981. In: *Models of Neural Networks*. Eds. Domany E., van Hemmen J.L., Schulten K. New York: Springer, 1994. P. 95-119.
48. Miltner W., Braun C., Arnold M., Witte H., Taub E. Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*. 1999. V. **397**. P. 434-436.
49. Mirollo R.E., Strogatz S.H. Synchronization of pulse-coupled biological oscillators. *SIAM J. Applied Mathematics*. 1990. V. **50**. P. 1645-1662.
50. Moran J., Desimone R. Selective attention gates visual processing in the extrastriate cortex. *Science*. 1985. V. **229**. P. 782-784.
51. *Neurobiology of Attention*. Eds. Itti G., Rees G., Tsotsos J.K. Amsterdam: Elsevier, 2005.
52. Opara R., Wörgötter F. Using visual latencies to improve image segmentation. *Neural Computation*. 1996. V. **8**. P. 1493-1520.
53. Opara R., Wörgötter F. A fast and robust cluster update algorithm for image segmentation in spine-lattice models without annealing - Visual latencies revisited. *Neural Computation*, 1998. V. **10**. P. 1547-1566.
54. Pal N.R., Pal S.K. A review on image segmentation techniques. *Pattern Recognition*. 1993. V. **26**. P. 1277-1294.
55. Palm G., Knoblauch A. Scene segmentation through synchronization. In: *Neurobiology of Attention*. Eds. Itti L., Rees G., and Tsotsos J.K. San Diego. CA: Elsevier, 2005. P. 618-623.
56. Pham D.L., Xu C., Prince J.L. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*. 2000. V. **2**. P. 315-337.
57. Polakowski W., Rogers S., Ruck D., Hoffmeister J., Raines R. Computer aided breast cancer detection and diagnosis of masses using difference of gaussians and derivative-based feature saliency. *IEEE Trans. Med. Imag.* 1997. V. **16**. P. 811-819.
58. Reynolds J., Desimone R. The role of neural mechanisms of attention in solving the binding problem. *Neuron*. 1999. V. **24**. P. 19-29.
59. Rhouma M.B.H., Frigui H. Self-organization of pulsed-coupled oscillators with application to clustering. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 2001. V. **23**. P. 180-195.

60. Rogers S. K., Colombi J. M., Martin C. E., Gainey J. C., Fielding K. H., Burns T. J., Ruck D. W., Kabrisky M., Oxley M. Neural networks for automatic target recognition. *IEEE Trans. Neural Networks*. 1995. V. **8**. P. 1153–1184.
61. Rutishauser U., Ross I.B., Mamelak A.N., Schuman E.M. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*. 2010. V. **464**. P. 903-907.
62. Sauseng P., Klimesch W., Gruber W.R., Birbaumer N. Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *Neuroimage*. 2008. V. **40**. P. 308-317.
63. Schillen T.B., König P. Binding by temporal structure in multiple feature domains of an oscillatory neural network. *Biol. Cybern.* 1994. V. **70**. P. 397-405.
64. Sederberg P.B., Kahana M.J., Howard M.W., Donner E.J., Madsen J.R. Theta and gamma oscillations during encoding predict subsequent recall. *J. Neurosci.* 2003. V. **23**. P. 10809-10814.
65. Shapiro L.G., Stockman G.C. *Computer Vision*. New Jersey: Prentice-Hall, 2001.
66. Sharon B., Galun M., Sharon D., Bash R., Brandt A. Hierarchy and adaptivity in segmenting visual scenes. *Nature*. 2006. V. **442**. P. 810-813.
67. Singer W. Neuronal synchrony: A versatile code for the definition of relations. *Neuron*. 1999. V. **24**. P. 49-65.
68. Singer W., Gray C.M. Visual feature integration and the temporal correlation hypothesis. *Ann. Rev. Neurosci.* 1995. V. **18**. P. 555-586.
69. Steinmetz P.N., Roy A., Fitzgerald P., Hsiao S.S., Johnson K.O., Niebur E. Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*. 2000. V. **404**. P. 187-190.
70. Strogatz S. *Nonlinear Dynamics and Chaos*. Perseus Publishing, 2000.
71. Ursino M., La Cara G.E. Modeling segmentation of a visual scene via neural oscillators: fragmentation, discovery of details and attention. *Network*. 2004. V. **15**. P. 69-89.
72. Ursino M., La Cara G.E., Sarti A. Binding and segmentation of multiple objects through neural oscillators inhibited by contour information. *Biol. Cybern.* 2003. V. **89**. P. 56-70.
73. Verschure P., König P. On the role of biophysical properties of cortical neurons in binding and segmentation of visual scenes. *Neural Computation*. 1999. V. **11**. P. 1113-1138.
74. Wang D.L. The time dimension for scene analysis. *IEEE Trans. Neural Networks*. 2005. V. **16**. P. 1401-1426.
75. Wang D.L., Terman D. Locally excitatory globally inhibitory oscillator network. *IEEE Trans. Neural Networks*. 1995. V. **6**. P. 283-286.
76. Wang D.L., Terman D. Image segmentation based on oscillatory correlation. *Neural Computation*. 1997. V. **9**. P. 805-836.
77. Weickert J. A review of nonlinear diffusion filtering. In: *Scale-space theory in computer vision*. Eds. Romany B., Florack L., Koenderick J., Viergever M. Berlin: Springer, 1997. P. 3-28.
78. Wikipedia-Segmentation (image processing).
URL: [http://en.wikipedia.org/wiki/Segmentation_\(image_processing\)](http://en.wikipedia.org/wiki/Segmentation_(image_processing)) (дата обращения: 28.06.2010).
79. Willshaw D., Buneman O., Longuet-Higgins H. Non-holographic associative memory. *Nature*. 1969. V. **222**. P. 960-962.
80. Womelsdorf T., Schoffelen J.M., Oostenveld R., Singer W., Desimone R., Engel A.K., Fries P. Modulation of neuronal interactions through neuronal synchronization. *Science*. 2007. V. **316**. P. 1609-1612.
81. Yu B., Zhang L. Pulse-coupled neural networks for contour and motion matching. *IEEE Trans. Neural Networks*. 2004. V. **15**. P. 1186-1201.

82. Zhang A. A survey on evaluation methods for image segmentation. *Pattern Recognition*. 1996. V. **29**. P. 1235-1346.
83. Zhao L., de Carvalho A., Li Z. Pixel clustering by adaptive pixel moving and chaotic synchronization. *IEEE Trans. on Neural Networks*. 2004. V. **15**. P. 1176- 1185.
84. Zhao L., Furukawa R.A., de Carvalho A. A network of coupled chaotic maps for adaptive multy-scale image segmentation. *Int. J. Neural Systems*. 2003. V. **13**. P. 129-137.
85. Zhao L., Macau E. A network of dynamically coupled chaotic maps for scene segmentation. *IEEE Trans. on Neural Networks*. 2001. V. **12**. P. 1375-1385.
86. Zhu S.C., Yuille Y. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1996. V. **18**. P. 884-900.
87. Абарбанель Г.Д.И., Рабинович М.И., Селверстон А., и др. Синхронизация в нейронных ансамблях. *Успехи физических наук*. 1996. Т. **166**. № 4. С. 363-390.
88. 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.
89. Вежневцев А., Баринаова О., Методы сегментации изображений: автоматическая сегментация. Графика и мультимедиа. URL: <http://cgm.computergraphics.ru/content/view/147> (дата обращения: 28.06.2010).
90. Гричук С., Кузьмина М.Г., Манькин Э.А. Селективная сегментация реальных изображений осцилляторно-сетевым методом. *Нейроинформатика-2007. IX Всероссийская научно-техническая конференция*, т. 3, Под ред. О.А. Мишулиной, Москва: МИФИ, 2007. С. 235-244.
91. Гричук С., Кузьмина М.Г., Манькин Э.А. Выделение объектов в зрительной сцене осцилляторно-сетевым методом. *Нейроинформатика-2009. XI Всероссийская научно-техническая конференция*, т. 3, Под ред.. О.А. Мишулиной, Москва: МИФИ, 2009. С. 167-172.
92. Казанович Я.Б., Шматченко В.В. Осцилляторные нейросетевые модели сегментации изображений и зрительного внимания. *Нейроинформатика-2004. VI Всероссийская научно-техническая конференция. Лекции по нейроинформатике*, т. 1. Под ред. Тюменцева Ю.В. Москва: МИФИ, 2004. С. 72-125.
93. Крюков В.И. Модель внимания и памяти, основанная на принципе доминанты и компараторной функции гиппокампа. *Журнал высшей нервной деятельности*. 2004. Т. **54**. № 1. С. 10-29.
94. Кузьмина М.Г., Манькин Э.А. Осцилляторная сеть для сегментации изображений: новые разработки. *Нейроинформатика-2004. VII Всероссийская научно-техническая конференция*, т. 1. Под ред. Мишулиной О.А. Москва: МИФИ, 2004. С. 262-268.
95. Чик Д., Борисюк Р.М., Казанович Я.Б. Режимы синхронизация в сети нейронов Ходжкина-Хаксли с центральным элементом. *Математическая биология и биоинформатика* (электронный журнал). 2008. Т. **3** № 1. С. 16-35. URL: [http://www.matbio.org/downloads/Chick2008\(3_16\).pdf](http://www.matbio.org/downloads/Chick2008(3_16).pdf) (дата обращения: 28.06.2010).

Материал поступил в редакцию 21.06.2010, опубликован 28.06.2010