

УДК 612.821

## Сегментация зрительных изображений: экспериментальные данные и моделирование

© 2021 г. **В. М. Бондарко**, доктор биол. наук; **М. В. Данилова**, канд. биол. наук;  
**В. Н. Чихман**, канд. техн. наук

Институт физиологии им. И.П. Павлова РАН, Санкт-Петербург

E-mail: [ymbond@gmail.com](mailto:ymbond@gmail.com)

Поступила в редакцию 16.09.2021

DOI:10.17586/1023-5086-2021-88-12-17-27

Экспериментальные данные по сегментации окружённых рамкой простых геометрических фигур впервые сопоставлены с результатами модели модулей, осуществляющей фильтрацию изображений в локальных участках поля зрения. В экспериментах показано, что изображения воспринимаются как два отдельных при достижении определённого расстояния между ними, зависящего от их размера и формы. При малых размерах расстояния согласованы с оптической функцией рассеяния и размером самых высокочастотных рецептивных полей нейронов первичной зрительной коры, при больших — с размерами модулей, оптимально описывающих изображения (с максимальным сохранением энергии у изображений при ограниченном числе фильтров). Изображения сегментируются, когда второе изображение (рамка) оказывается за пределами модуля. Результаты подтверждены ранее полученными нами данными по изучению иллюзии Оппель–Кундта и оценке ширины пространственных интервалов.

**Ключевые слова:** сегментация, моделирование, пространственно-частотный анализ, модель модулей, обрамление картин, «Чёрный квадрат» К. Малевича.

**Коды OCIS:** 330.7326 330.4060 330.5510.

### ВВЕДЕНИЕ

Сегментация зрительных сцен — один из ключевых аспектов зрительного восприятия, а различные методы сегментации широко используются в работе искусственных нейронных сетей [1–3]. Наиболее изученной является сегментация текстур. Максимальный вклад в описание механизмов зрительного восприятия по различению текстур, по мнению многих исследователей, внесли Юлеш и Бэк. Они показали, что различение текстур человеком происходит мгновенно, ещё до момента фокусировки внимания на объекты (preattentive discrimination) при различии в таких свойствах, как освещённость, цвет, размер элементов или наклон отрезков линий, образующих текстуру [4–6]. В результате многочисленных

исследований Юлешем [7, 8] был сделан вывод о том, что мгновенное различение текстур зависит только от локальных признаков образующих их элементов (текстонов).

Любые локальные признаки могут быть записаны на языке спектральных характеристик. Спектральные модели, предлагаемые для различения текстур [9–16], являются частным видом моделей локального частотного анализа с использованием фильтров, имеющих различную ориентационную и частотную настройку. Завершающим этапом различения текстур является сегментация сцены и выделение фигуры из фона.

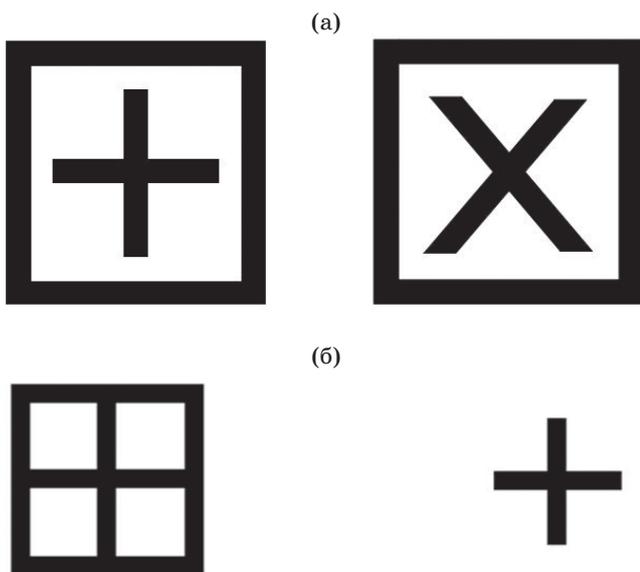
В зрительном восприятии размеры локальных окон при сегментации Юлеш ассоциировал с избирательным (фокальным) вниманием [7].

Его представления на языке спектрального анализа с учётом модульной организации зрительной системы [17] сводятся к подбору размера оптимального модуля для выделения объекта из фона и его описания. Сканирование зрительного поля локальным окном осуществляет механизм внимания. Каким образом происходит подбор оптимального модуля, какое соотношение между размерами локального и глобального окон должно выполняться, до сих пор остаётся малоизученным вопросом.

Ранее в психофизических исследованиях мы показали, что сегментация объектов зависит как от размера, так и от формы изображений [18, 19]. Более того, она связана с оценкой размера объектов [20]. В настоящей работе представлены новые экспериментальные данные, полученные на большой группе наблюдателей в одинаковых условиях предъявления стимулов, и осуществлено моделирование результатов по сегментации простых геометрических фигур — косых и прямых крестов, заключённых в рамки (рис. 1а). Выбор этих изображений не случаен. Косые кресты в рамках Аткинсон [21] использовала в качестве стимулов при изучении краудинг-эффекта (явления ухудшения восприятия объектов при близком расположении окружения) на пределе разрешения зрительной системы. Краудинг-

эффект на пределе разрешения был успешно нами промоделирован [22] с применением пространственно-частотных (ПЧ) фильтров, которыми являются наименьшие по размеру самые высокочастотные рецептивные поля (РП) зрительной коры. Показано, что ухудшение опознания объекта происходит при попадании окружающих его изображений в тормозные зоны РП нейронов, оптимально реагирующих на объект. Размеры возбуждающих и тормозных зон самых высокочастотных РП согласованы с оптической функцией рассеяния и мозаикой фоторецепторов на сетчатке [23]. В настоящей работе отобраны изображения крестов в рамках, поскольку их можно использовать, как при исследовании краудинг-эффекта, так и сегментации.

В качестве модели нами рассмотрена одна из разновидностей модели модулей, предложенная В.Д. Глезером [24] на основании полученных электрофизиологических данных при исследовании РП нейронов стриарной коры. В модели осуществляется разложение изображений в конечные обобщённые ряды Фурье с использованием базисных функций таких, как  $\sin$  и  $\cos$  [25] или элементы Габора [26]. В рамках электрофизиологических исследований трудно определить точный вид базисных функций из-за их схожести, поэтому оба вида функций допустимы. Мы выбрали в качестве базисных  $\sin$  и  $\cos$ , поскольку при разложении в ряд Фурье по определению они заданы на конечном промежутке и нам требовалось оценить размеры окон. Элементы Габора заданы на всей плоскости, при определении размера локального окна пришлось бы отсекал какой-то уровень на гауссовой кривой, что делало бы эту процедуру довольно произвольной. Следует отметить, что в радиотехнике и оптике спектральный анализ, в частности, разложение в ряды Фурье используют для периодических колебательных сигналов. Такая процедура корректна и для любой кусочно-непрерывной функции на конечном промежутке разложения [27–29], которую можно при желании периодически продолжить на всю числовую ось. Однако для нашего анализа этого не требуется. Разложение в ряды Фурье представляется физиологичным в связи с конечными размерами РП и выявленной дискретностью распределения пространственных частот, на которые настроены РП нейронов [30].



**Рис. 1.** Используемые в исследовании изображения. (а) — примеры тестовых изображений, (б) — ответная карта наблюдателя для случая прямых крестов.

Цель настоящей работы заключается в описании нейрофизиологических механизмов, обеспечивающих целостное восприятие простых геометрических объектов и их сегментацию посредством сопоставления экспериментальных данных и моделирования. Изучение сегментации геометрических объектов разных размеров (от малых на пределе разрешения до относительно больших) с последующим моделированием позволяет сравнить пространственные характеристики изображений в двух зрительных задачах и высказать предположения относительно механизма сегментации, а также определить размеры локальных окон.

Необходимо подчеркнуть, что в настоящее время для сегментации и классификации изображений применяют свёрточные нейронные сети с фильтрами разного размера и настройки [31–33], что обеспечивает инвариантность при работе с изображениями разных масштабов. При обучении нейронных сетей используют комбинированные методы локального и глобального анализа [34–37]. Их совместное применение может ускорить процесс поиска цели и сегментацию. В ряде работ быстрое действие обеспечивается применением алгоритмов на основе принципа наименьшего действия, ключевого принципа для обучения сетей и обеспечения их эффективной работы [38]. Полученные в настоящем исследовании результаты могут способствовать дальнейшему усовершенствованию работы нейронных сетей.

## МЕТОДИКА

**Наблюдатели.** В экспериментах приняли участие 11 наблюдателей в возрасте от 18 до 52 лет с нормальным или откорректированным зрением. Исследование было одобрено Этическим комитетом СПбГУ. Наблюдатели дали письменное согласие на проведение исследований.

**Стимулы.** Стимулами служили изображения крестов, вокруг которых нарисованы квадратные рамки. Размер крестов определялся как сторона квадрата, в который они были вписаны. Так как нас интересовала зависимость процесса сегментации и от формы выбранных объектов, то помимо косых крестов, использованных ранее при изучении краудинг-эффекта [21], также использовались прямые кресты (рис. 1а).

Ширина линий рамки составляла  $1/10$  от размера рамки. Линии у крестов имели ту же ширину, а размеры их варьировали. На рис. 1а показаны примеры применявшихся в исследованиях стимулов (слева прямые кресты, справа косые кресты).

**Процедура.** Стимулы предъявлялись в случайном порядке в центре экрана дисплея. В каждой серии эксперимента фиксировался размер рамки, а размеры крестов менялись. Для каждого размера рамки было взято по 20 размеров крестов: от полного соприкосновения с рамкой до  $1/3$  размера рамки. Шаг изменения размеров крестов был пропорционален размеру рамки. Размеры рамок имели следующие значения: 11, 13, 14,5, 18, 21, 25, 36, 71, 107, 142 и 180 угл. мин. Размеры крестов в серии менялись случайным образом. Изображения косого и прямого крестов предъявлялись по очереди. В одном эксперименте изображение с каждым размером креста при фиксированном размере рамки предъявлялось 20 раз.

Испытуемому была дана ответная карта, на которой были нарисованы четыре изображения. На рис. 1б показаны два из них, соответствующие ответной карте для прямых крестов, для косых крестов изображения на ответной карте были аналогичными. Наблюдатель должен был ответить на вопрос — на что больше похож предъявляемый стимул: на крест, слившийся с рамкой, или на изолированный крест, и нажать соответствующую клавишу на клавиатуре. Каждый наблюдатель принял участие не менее, чем в двух экспериментах с одним и тем же размером рамки с перерывами в экспериментах более двух недель для контроля за стабильностью выбранного критерия разделения изображений.

**Аппаратура.** Стимулы предъявлялись наблюдателям на мониторе LG Flatron 795FT Plus с размером диагонали 17” при разрешении экрана  $1024 \times 768$  пикселей с частотой вертикальной развёртки 100 Гц. Яркость экрана была  $80 \text{ кд/м}^2$ , изображений  $5 \text{ кд/м}^2$ . Время экспозиции изображений не было ограничено. Стимулы предъявлялись бинокулярно с расстояния 3 м до экрана. Программное обеспечение эксперимента реализовано в среде Delphi 6.

**Обработка данных.** Для определения размеров, при которых меняется ответ («крест» — «рамка») был использован пробит-анализ.

Были построены психометрические функции для ответов «стимул похож на рамку» по 40 предъявлениям каждого размера крестов при каждом фиксированном значении размера рамки для каждого наблюдателя. С помощью метода наименьших квадратов психометрические функции были приближены к функциям нормального распределения. Разницы между экспериментальными данными и полученными нормальными распределениями оценены по методу  $\chi^2$ . Величины средних значений нормальных распределений соответствуют тем критическим максимальным размерам крестов, при которых наблюдатели отделяют их от рамок. Это тот размер, когда испытуемые с одинаковой вероятностью говорят, что объект похож на отдельный крест и на крест в рамке.

**Моделирование.** В модели модулей изображения разлагаются в двумерные конечные ряды Фурье с помощью базисных функций  $\sin$  и  $\cos$  на квадратных участках, больших, чем изображения, или равных по размеру. Размеры участков варьируются. Для каждого изображения подбирается размер модуля таким образом, чтобы сохранить наибольшее количество энергии у изображения, равное сумме квадратов коэффициентов ряда Фурье. Такой модуль назван оптимальным, так как при его использовании происходит наименьшая потеря информации об изображении. Ранее для других изображений было показано [39], что каждое изображение наилучшим образом описывается модулем своего размера. Мы полагаем, что размеры локальных окон определяются областями, на которых функционируют оптимально описывающие дан-

ные конкретные изображения РП нейронов. Для подтверждения справедливости этого предположения сопоставляются экспериментально полученные и рассчитанные в моделях размеры окон.

Для изображений больших размеров мы попробовали подобрать размер оптимальных модулей с конечным числом гармоник, поскольку только при довольно больших размерах при естественном ограничении на высокую пространственную частоту из-за физических характеристик оптики глаза могут функционировать фильтры с большим числом гармоник. Были взяты модули с 3, 4 или 5 гармониками.

## РЕЗУЛЬТАТЫ

Психометрические функции у девяти из одиннадцати наблюдателей имели правильную форму с высокой точностью приближения к нормальным распределениям ( $p < 0,05$ ). Критерий перехода ответа от «креста» к «рамке» у каждого испытуемого практически не менялся от опыта к опыту. Чётко прослеживалась постоянная тенденция: смена ответов происходила при меньших размерах косых крестов, чем для прямых крестов. У двух оставшихся наблюдателей критерий разделения изображений на крест и рамку был другой. Они отделяли рамку от крестов при условии чётко видимого разрыва между краями изображений, то есть при расстоянии около 1 угл. мин между краями объектов. Их данные не взяты в рассмотрение для дальнейшего анализа.

В таблице приведены усреднённые по данным девяти наблюдателей значения разме-

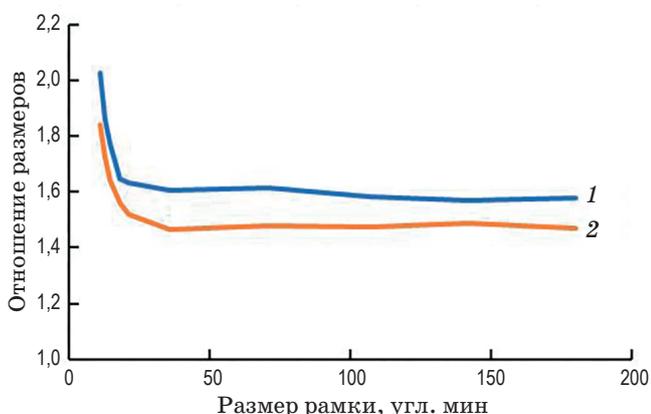
### Максимальные размеры крестов и расстояния до рамок, при которых происходит сегментация прямых (ПК) и косых (КК) крестов

Изображения	Размеры рамок (угл. мин)									
	11	13	14,5	18	21	36	71	107	142	180
<i>Размеры крестов, при которых происходит сегментация (угл. мин)</i>										
КК	5,4	7,0	8,2	10,9	12,8	22,4	44,0	67,5	90,4	113,9
ПК	6,0	7,5	8,8	11,5	13,8	24,6	48,0	72,5	95,3	122,4
<i>Расстояния до рамок (угл. мин)</i>										
КК	2,78	3,01	3,17	3,53	4,08	6,79	13,52	19,75	25,78	33,04
ПК	2,51	2,73	2,83	3,23	3,59	5,71	11,51	17,23	23,35	28,78

ров крестов, при которых происходила смена ответа «крест» на ответ «рамка» отдельно для прямых (ПК) и косых (КК) крестов. Соседние значения для разных размеров рамок достоверно отличаются друг от друга по критерию Стьюдента ( $p < 0,01$ ). Достоверно отличаются и размеры для косых и прямых крестов при одинаковых размерах рамки ( $p < 0,05$ ). В этой же таблице показаны расстояния между крестами и рамками, при которых происходит сегментация стимулов. Эти расстояния получены посредством вычитания из размера рамки размера креста с последующим делением пополам.

Во всех случаях размеры косых крестов, измеренные как сторона квадрата, в который они вписаны, меньше размеров прямых крестов. Есть отличия в размерах, если их оценивать и как диаметр описанной вокруг изображений окружности. При равенстве длин линий у крестов сторона квадрата для косого креста должна быть примерно в  $1,4 (\sqrt{2})$  раза меньше, что также не выполняется. Соответственно расстояния до рамок для косых крестов больше. Необходимо отметить, что расстояния до рамок при малых размерах стимулов соответствуют размерам критических зон в краудинг-эффекте на пределе разрешения зрительной системы [22], что может свидетельствовать о связи механизмов сегментации и краудинг-эффекта.

На рис. 2 показаны отношения усредненных по данным девяти испытуемых размеров рамок к размерам прямых и косых крестов, полученные при условии их делимости.

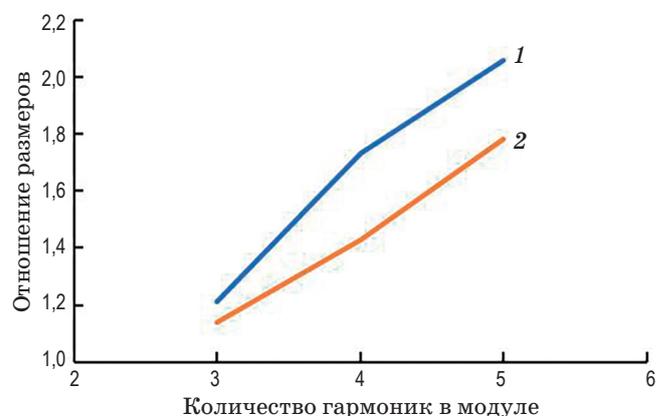


**Рис. 2.** Отношение размеров рамок к размерам косых (1) и прямых (2) крестов при условии их сегментации в зависимости от размера рамки.

Из приведённых графиков видно, что при увеличении размеров рамок отношения размеров рамок к размерам крестов сначала уменьшаются, а затем перестают изменяться. Стабилизация отношений наступает при размерах рамок, превышающих 21 угл. мин. Для наклонных крестов у всех наблюдателей отношения размеров имеют большие значения, когда кресты начинают восприниматься отдельно от рамок. Это означает, что косые кресты кажутся изолированными от рамок при меньшем физическом размере, чем прямые кресты, т.е. требуется большее расстояние между рамкой и объектом.

Результаты моделирования приведены на рис. 3, где показаны отношения размеров оптимальных модулей (см. методику) к размерам изображений прямого и косого крестов.

Видно, что размеры модулей растут с добавлением числа гармоник. Для косых крестов модули имеют большие размеры, чем для прямых крестов, что соответствует экспериментальным данным. Для модуля с четырьмя гармониками отношения размеров модуля к размерам косых и прямых крестов близки к экспериментальным данным, полученным для больших размеров стимулов. Таким образом, мы привели результаты моделирования, свидетельствующие в пользу предложенной нами гипотезы: успешная сегментация изображений происходит при выходе рамок за пределы областей, на которых заданы фильтры, обрабатывающие стимулы.



**Рис. 3.** Отношения размеров оптимальных модулей, на которых заданы базисные функции при разложении косых (1) и прямых (2) крестов в двумерные ряды Фурье, к размерам изображений в зависимости от числа гармоник в модуле.

## ОБСУЖДЕНИЕ

Субъективная методика сегментации вложенных друг в друга простых геометрических фигур позволила определить размеры, при которых происходит разделение стимулов на отдельные составляющие для выбранной группы объектов. Анализ психометрических функций показал низкий разброс данных у группы наблюдателей из девяти человек (переход от восприятия фигуры как целостной к восприятию вложенного креста как отдельного объекта происходил в пределах двух-четырёх шагов изменения размеров крестов), что позволяет предположить объективный характер субъективных оценок.

При увеличении размеров уменьшается отношение размера целого изображения к размеру креста от 1,8 до 1,5 (прямой крест) или от 2 до 1,6 (косой крест), при меньших значениях отношений наблюдается стабилизация этой функции. Эти отношения достоверно отличаются: при разделении косых крестов от рамок эти стимулы имеют меньший размер, чем прямые кресты. Стабилизация отношений размеров полного изображения (рамка + крест) к размерам крестов происходит при увеличении величины изображения до 20–25 угл. мин. Эта величина примерно равна размеру центрального островка фовеолы, равного 20–30 угл. мин у разных наблюдателей и подтверждает гипотезу о специфической роли центральной части фовеолы и связанных с ней РП в процессе сегментации и объединении изображений [19, 40–42]. Инвариантное к размеру описание наблюдается при размерах изображений, равных или превышающих размер фовеолы, и определяется совместным функционированием нейронов с РП, настроенными на разные частоты.

**Сегментация и пространственно-частотный анализ.** В нашей работе показано, что для описания объектов с минимальной потерей информации конечным набором гармоник (было взято 3, 4 и 5 гармоник) размер оптимального модуля превосходит объект в 1,2–2,1 раз в зависимости от формы этого объекта и от количества гармоник. Аналогичные размеры были получены [39] для другого набора стимулов — стилизованных букв греческого алфавита. Оптимальный средний для всех букв модуль имел размер в 1,5 раз превосходящий размер стимулов. Этот размер совпадает

с полученными экспериментальными данными по сегментации крестов и при моделировании модулем с четырьмя гармониками в данной работе. Выбор модуля, превосходящего по размеру внутренний объект, демонстрирует тот факт, что окружение однородным полем необходимо для более точного описания объекта конечным набором фильтров. То есть, в данном случае подбор окна основан на оптимальном описании изображений крестов набором фильтров, а попадание соседнего изображения в зону их функционирования приводит к ухудшению описания тестового изображения.

Применение модели модулей в нашем случае следует рассматривать как самое первое приближение к описанию процесса сегментации простых геометрических фигур. Для этого класса объектов можно сказать, что процесс сегментации включает оптимальное описание ПЧ фильтрами с наименьшей потерей информации об объекте при достижении пространственной изоляции. Модель модулей представляет собой одну из предпринятых во многих работах попыток построения нейронных сетей с учётом взаимодействия ПЧ фильтров, в роли которых выступают РП нейронов первичной зрительной коры. По своей сути она близка к иерархическим моделям организации зрительной системы. Эти модели (или модели пирамид) впервые были предложены для решения чисто технической задачи по уменьшению избыточности в передаче информации об изображениях [43]. Под ними подразумевается последовательная иерархическая обработка изображений элементами-фильтрами, осуществляемая послойно. Элементы более высокого уровня образованы путём объединения элементов предыдущего уровня. Они подобны элементам предыдущего уровня и отличаются только размером: на каждом последующем уровне размеры элементов увеличиваются пропорционально на какую-то определённую величину, характерную для заданной пирамиды. Каждый уровень, в частности, может обрабатывать изображения посредством свёртки с разностью двух гауссиан определённого размера. В моделях возможно и применение фильтров, имеющих какую-либо ориентацию [44]. Подобные операции широко используются при обработке изображений в искусственных нейронных сетях. Ф. Кемпбелл

и Ю.Е. Шелепин [40] предположили, что центральный островок фовеолы (около 30 угл. мин.) и связанные с ним ретиноптически РП нейронов разных областей могут играть роль минимального по размерам слоя пирамиды. О фовеоле как о низшем слое зрительной пирамиды свидетельствуют как многочисленные данные по инвариантному к размеру восприятию изображений, так и результаты по опознанию, различению, обнаружению и чёткому видению изображений [19, 45]. Вилсоном и др. [46] в психофизических исследованиях была выявлена дискретность в настройке частотных каналов: наиболее чувствительный настроен на 4 цикл/град. с размером фильтра около 20 угл. мин, самый высокочастотный на 16 цикл/град. В исследованиях по опознанию прямоугольных решёток [45] показано, что размер высокочастотных фильтров составляет около 0,4 угл. град., что равно 24 угл. мин. То есть ограничение в 4 гармоник становится естественным для центральной части фовеолы. При этом размеры зон самого высокочастотного канала согласованы с мозаикой рецепторов на сетчатке и оптической функцией рассеяния [23].

Прослеживается интересная аналогия, которую нельзя отнести к случайной. В усовершенствованном варианте модели модулей [26] предпринята попытка обеспечить инвариантность описания изображений к размеру. В нейрофизиологических исследованиях В.Д. Глезера и др. [30] была показана дискретность в распределении размеров РП. На основании этих данных модули также были построены с дискретными размерами, что привело к отсутствию инвариантности к размеру в описании изображений. На основании экспериментально полученного распределения у модулей были выбраны размеры с разницей в полкватвы, что составляло величины 1,0; 1,41; 2,0; 2,83 и т.д. Моделирование в настоящем исследовании показало, что размер оптимального модуля соответствует размеру последующего модуля в этой цепочке. То есть, получается, что размеры модулей, выявленные в электрофизиологических экспериментах, связаны с оптимальным описанием объектов. Таким образом, в настоящей работе получено соответствие между данными нейрофизиологических, психофизических и модельных исследований. Возможно, что использование

искусственных расширенных свёрточных нейронных сетей с ядрами подобного размера и шага окажется эффективным при сегментации и классификации изображений.

*Сегментация и краудинг-эффект.* О связи между сегментацией и размерами РП нейронов свидетельствуют данные, полученные при минимальных размерах рамки, поскольку в этом случае расстояния до рамок соответствуют размерам критических зон в краудинг-эффекте на пределе разрешения зрительной системы [47, 48], в пределах которых опознание формы объекта ухудшается. Размер зон краудинг-эффекта, в свою очередь, совпадает с размером тормозных зон наименьших по размеру высокочастотных РП [22]: на пределе разрешения ухудшение восприятия происходит при попадании добавочных изображений в тормозные зоны РП нейронов, обрабатывающих весь стимул. Следовательно, сегментация малых по размеру изображений происходит тогда, когда окружение теста не попадает в описывающие его высокочастотные РП, что подтверждает наши представления о сегментации моделью модулей.

*Сегментация и супрематизм.* С проблемой сегментации связано, на наш взгляд, создание Казимиром Малевичем стиля супрематизм. Его картины, особенно «Чёрный квадрат», вызывают споры и интерес до сих пор, несмотря на использование простейших форм или абстрагирования от формы как таковой.

В различных вариантах «Чёрного квадрата» меняются соотношения размеров целого холста, окантованного в раму, к стороне самого квадрата: у двух картин соотношение равно 1,27, у двух — 1,41. Казимир Малевич менял соотношение размеров полотна и самого квадрата интуитивно, для достижения большего эффекта. (Необходимо отметить, что проблеме обрамления картин и выбора фона в живописи и графике художники уделяли большое внимание ещё со времен Возрождения.) Самые знаменитые — поздняя и ранняя картины — написаны в 1915 и 1929 и находятся в Третьяковской галерее. Картины одинаковы по размеру, отличаются палитрой и текстурой квадрата. У них размер рамы превосходит размер квадрата в 1,41 раза, т.е. тёмный квадрат окружен светлым полотном. Близкая величина получена нами при моделировании моделью модулей изображений различной формы

при оценке размера оптимального модуля, что может свидетельствовать в пользу нашего предположения о процедуре подбора локального окна. (Напомним, что в модифицированной модели модулей [26] эти размеры квадрата и рамки соответствуют размерам двух соседних модулей.) Проведение экспериментов по сегментации квадратов и рамок затруднено из-за отсутствия адекватной карты ответа.

**Сегментация и оптические иллюзии.** В нашем предыдущем исследовании [20] было показано, что сегментация может быть связана с оценкой размера: изображения кажутся равными, когда их описывают фильтры, заданные на промежутках одного размера. Близкие к квадратам по форме изображения — это пространственные интервалы — промежутки пространства, ограниченные двумя параллельными отрезками линий (рис. 4б). С оценкой их размера связано несколько зрительных иллюзий.

При оценке размера наблюдается наибольшее количество оптических иллюзий, которые возникают из-за наличия контекста/окружения. Широко известна иллюзия Опфель–Кундта (ОК), в которой размер пространственного интервала переоценивается, если внутри него на равном расстоянии располагаются аналогичные отрезки линий (рис. 4б). С увеличением числа линий при одинаковой ширине интервала сначала иллюзия возрастает, а затем стабилизируется при количестве линий от 5 до 11 [49, 50]. Стабилизация происходит при переоценке ширины интервала в 1,15–1,3 раза. Эта величина сопоставима

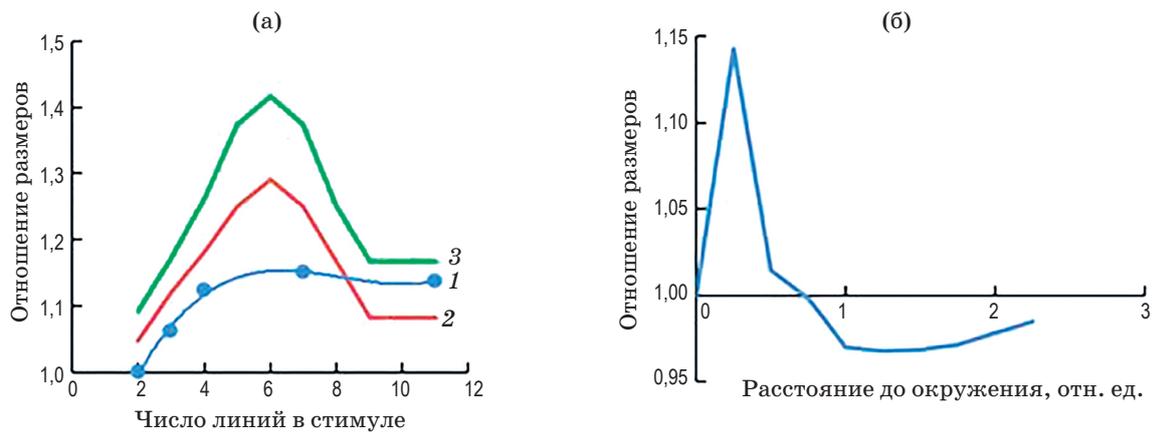


**Рис. 4.** Иллюстрации к обсуждению. (а) — изображение «Чёрного квадрата» Казимира Малевича из Третьяковской галереи. (б) — иллюзия Опфель–Кундта. Ширина пространственного интервала между двумя левыми крайними линиями равна ширине промежутка между первой и восьмой линиями справа, но промежуток справа кажется большим по размеру.

с отношением размеров рамок и квадратов для картин Малевича.

Инвариантная по отношению к размеру интервала экспериментальная зависимость от числа линий в иллюзии ОК аппроксимирована нами ранее [50] также моделью модулей. Сравнение результатов моделирования и экспериментальных данных приведено на рис. 5а. Отношение размеров модулей с четырьмя и пятью гармониками (кривые 2 и 3) к размеру изображений примерно соответствует экспериментально измеренной переоценке размеров интервалов (кривая 1) при стабилизации этой функции. Можно заключить, что иллюзия ОК связана с обработкой изображений ПЧ фильтрами и пространственным взаимодействием между ПЧ каналами, а её размер определяется размерами модулей.

Дополнительные линии, расположенные снаружи интервала при малых расстояниях, также приводят к переоценке его ширины [51]. Незначительная недооценка или отсутствие иллюзии наблюдается при увеличении расстояния. (Аналогичная картина наблюдается и при образовании интервалов элементами Габора [52].) Максимальная переоценка возникает при наименьшем использованном в исследовании расстоянии между дополнительными линиями и интервалом, равном четверти его ширины (рис. 5б). Влияние дополнительных контуров можно объяснить как их ассимиляцией с линиями, задающими интервал, так и взаимодействием между ПЧ каналами, образующими нейронную сеть. Максимальная переоценка интервала соответствует величине иллюзии ОК при её стабилизации (ср. рис. 5а и 5б). Для максимального использованного в исследовании [51] размера интервала (33,4 угл. мин) иллюзия исчезает у большинства наблюдателей при расстоянии, равном половине интервала (рис. 5б): при таких расстояниях окружение оказывает минимальное воздействие на оценку размера. Отношение расстояний между линиями, на которых наблюдается максимальная иллюзия, к ширине интервала равно 1,5, что примерно соответствует отношению размера рамки к длине стороны квадрата в «Чёрном квадрате». Таким образом, картина «Чёрный квадрат» учитывает оптимальную сегментацию благодаря наличию самого квадрата и оставленного светлого полотна вокруг него.



**Рис. 5.** Сравнение размеров. (а) — преувеличение размера пространственных интервалов, полученное в иллюзии Опфель–Кундта экспериментально (1) и в модели модулей с четырьмя (2) и пятью гармониками (3) в зависимости от числа линий в стимуле. (б) — искажение оценки размера пространственного интервала в зависимости от расстояния до двух боковых линий. Расстояния выражены в долях от размера интервала. На (а) и (б) приведены отношения кажущихся размеров к физическим размерам стимулов.

Наблюдается аналогия между геометрическими иллюзиями, связанными с оценкой ширины пространственных интервалов, и сегментацией простых объектов: расстояния между объектами при исчезновении иллюзии соответствуют расстояниям, необходимым для успешной сегментации, а стабилизация сегментации и иллюзии ОК в модели происходит при размерах модуля, превосходящих размер стимула. Эти результаты хорошо согласуются с чисто интуитивным представлением о том, что для того, чтобы восприятию объекта не мешали другие объекты, его надо окружить пустым полем. При этом размеры пустого поля довольно большие и определяются не только оптическими факторами. С точки зрения моделирования при использовании ПЧ анализа можно говорить о том, что размер этого поля определяется как оптимальный размер окна, обеспечивающий сохранение максимального количества информации об объекте.

## ЗАКЛЮЧЕНИЕ

Исследования по сегментации вложенных чёрно-белых изображений показали зависимость этого процесса не только от формы, но и от размера простых геометрических объектов — крестов, помещённых в центр квадратной рамки той же толщины. Впервые продемонстрирована инвариантность этого процесса к размеру у большой группы наблюдателей

при размерах, превосходящих размер фoveолы (около 20–30 угл. мин). При малых размерах изображений на большой группе наблюдателей подтверждены ранее полученные нами данные [18] у двух наблюдателей о связи процесса сегментации с краудинг-эффектом на пределе разрешения зрительной системы, обусловленным функционированием самых высокочастотных РП первичной зрительной коры. Впервые показано, что оптимальное описание изображений с наименьшей потерей информации конечным набором фильтров, осуществленное нами посредством моделирования, связано с сегментацией объектов. Можно предположить, что для успешной сегментации изображений, обрабатываемых конечным набором ПЧ фильтров, необходим выход другого изображения за пределы области, на которых эти фильтры заданы. Впервые выявлена согласованность между данными нейрофизиологических, психофизических и модельных исследований по отношению к модели модулей: дискретность модулей связана с оптимальным описанием изображений. Скоординированное сочетание использования методов локального и глобального Фурье-анализа может быть использовано в искусственных нейронных сетях для успешного выделения простых геометрических объектов из окружения и их сегментации при наименьшей потере информации об объекте. Построение искусственных расширенных свёрточных

нейронных сетей с ядрами, отличающимися по размеру на полоктавы, представляется эффективным как для сегментации, так и классификации изображений.

Работа выполнена при поддержке Госпрограммы 47 ГП «Научно-технологическое развитие Российской Федерации» (2019–2030), тема 0134-2019-0005.

## ЛИТЕРАТУРА

1. Луцив В.Р. Сверточные искусственные нейронные сети глубокого обучения // Оптический журнал. 2015. Т. 82. № 8. С. 11–23.
2. Цыцулин А.К., Бобровский А.И., Морозов А.В., Павлов В.А., Галева М.А. Применение свёрточных нейронных сетей для автоматической селекции малоразмерных искусственных космических объектов на оптических изображениях звёздного неба // Оптический журнал. 2019. Т. 86. № 10. С. 30–38.
3. Xie Y., Zhu F., Fu Y. Main-secondary network for defect segmentation of textured surface images // IEEE Winter Conference on Applications of Computer Vision (WACV). 2020. P. 3531–3540.
4. Beck J. Textural segmentation, second-order statistics, and textural elements // Biological Cybernetics. 1983. V. 48. P. 125–130.
5. Julesz B. A theory of preattentive texture discrimination based on first-order statistics of textons // Biological Cybernetics. 1981. V. 41. P. 131–138.
6. Julesz B. Textons, the elements of texture perception, and their interactions // Nature. 1981. V. 290. P. 91–97.
7. Julesz B. Towards an axiomatic theory of preattentive vision. Dynamic aspects of neocortical function. Toronto, Ontario, Canada: Wiley, 1984. P. 585–612.
8. Julesz B. Texton gradients: The texton theory revisited // Biological Cybernetics. 1986. V. 54. P. 245–251.
9. Fogel I., Sagi D. Gabor filters as texture discriminator // Biological Cybernetics. 1989. V. 61. P. 103–113.
10. Landy M.S., Bergen J.R. Texture segregation and orientation gradient // Vision Research. 1991. V. 31. P. 679–691.
11. Beck J., Sutter A., Ivry R. Spatial frequency channels and perceptual grouping in texture perception // Computer Vision, Graphics and Image Processing. 1987. V. 37. P. 299–325.
12. Bovik A.C., Clark M., Geisler W.S. Multichannel texture analysis using localized spatial filters // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1990. V. 12. P. 55–73.
13. Malik J., Perona P. Preattentive texture discrimination with early vision mechanisms // J. Opt. Soc. Am. A. 1990. V. 7. P. 923–932.
14. Turner M.R. Texture discrimination by Gabor functions // Biological Cybernetics. 1986. V. 55. P. 71–82.
15. Arsenaault E., Yoonessi A., Baker C. Higher order texture statistics impair contrast boundary segmentation // J. Vision. 2011. V. 11(10):14. P. 1–15.
16. Prakash M.J., Kezia S., I. Santhi Prabha S., Kumar V.V. A new approach for texture segmentation using gray level textons // International J. Signal Processing, Image Processing and Pattern Recognition. 2013. V. 6. № 3. P. 81–90.
17. Blasdel G., Salama G. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex // Nature. 1986. V. 321. P. 579–585.
18. Бондарко В.М., Данилова М.В. Оценка целостности зрительных объектов в зависимости от их размеров // Физиология чел. 1996. Т. 22. № 3. С. 65–70.
19. Шелепин Ю.Е., Бондарко В.М., Данилова М.В. Конструкция фовеолы и модель пирамидальной организации зрительной системы // Сенсорные системы. 1995. Т. 9. № 1. С. 86–96.
20. Бондарко В.М., Данилова М.В. Оценка размеров зрительных объектов, имеющих различную форму, зависит от расстояния между объектами и условий их предъявления // Сенсорные системы. 1997. Т. 11. № 3. С. 300–311.
21. Atkinson J. Review of human visual development: crowding and dyslexia // Vision and visual dysfunction (ed.J.R.Cronly- Dillon). 1991. V. 13. Vision and visual dislexia (ed.J.F.Stein). P. 44–57.
22. Danilova M.V., Bondarko V.M. Foveal contour interactions and crowding effects at the resolution limit of the visual system // J. Vision. 2007. V. 7(2). № 25. P. 1–18.
23. Shelepin Y.E., Bondarko V.M. Resolving ability and image discretization in the visual system // Neurosci. Behav. Physiol. 2004. V. 34. P. 147–157.
24. Глезер В.Д. Зрение и мышление. Л.: Наука, 1985. 300 с.

25. Вол И.А. Пространственно-частотная модель гиперостроты зрительной системы // Сенсорные системы. 1988. Т. 2. № 2. С. 133–138.
26. Kaliteevsky N.A., Semenov V.E., Glezer V.D., Gauselman V.E. Algorithm of invariant image description by the use of a modified Gabor transform // Applied optics. 1994. V. 33. № 23. P. 5256–5261.
27. Алюков С.В. Аппроксимация ступенчатых функций в задачах математического моделирования // Матем. моделирование. 2011. Т. 23. № 3. С. 75–88.
28. Рудин У. Основы математического анализа. М.: Мир, 1976. 320 с.
29. Шилов Г.Е. Математический анализ. Специальный курс. М.: ГИФМЛ, 1961. 436 с.
30. Glezer V.D., Yakovlev V.V., Gauselman V.E. Harmonic basis function for spatial coding in the cat striate cortex // Visual Neurosci. 1989. V. 3. P. 351–383.
31. Жукова О.В., Малахова Е.Ю., Шелепин Ю.Е. Джоконда и неопределенность распознавания улыбки человеком и искусственной нейронной сетью // Оптический журнал. 2019. Т. 86. № 11. С. 40–50.
32. Малахова Е.Ю. Пространство описания зрительной сцены в искусственных и биологических нейронных сетях // Оптический журнал. 2020. Т. 87. № 10. С. 50–58.
33. Титаренко М.А., Малашин Р.О. Метод улучшения изображений с помощью глубоких нейронных сетей при использовании высокоуровневой информации // Оптический журнал. 2020. Т. 87. № 10. С. 59–68.
34. Жукова О.В., Васильев П.П. Перестройка нейронной сети и изменение стратегий операторов в процессе распознавания изображений лиц // Оптический журнал. 2020. Т. 87. № 10. С. 25–37.
35. Хараузов А.К., Шелепин Ю.Е., Цветков О.В., Жукова О.В., Пронин С.В. Методы маскировки угрожающих изображений и электрофизиологические маркеры их неосознанного восприятия // Оптический журнал. 2020. Т. 87. № 10. С. 69–80.
36. Шелепин Ю.Е., Хараузов А.К., Жукова О.В., Пронин С.В., Куприянов М.С., Цветков О.В. Маскировка и обнаружение скрытых сигналов в динамических изображениях // Оптический журнал. 2020. Т. 87. № 10. С. 89–102.
37. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // 18<sup>th</sup> International Conference / MICCAI 2015. Part III. Munich, Germany. October 5–9. 2015. P. 234–241.
38. Малашин Р.О. Принцип наименьшего действия в динамически конфигурируемых системах анализа изображений // Оптический журнал. 2019. Т. 86. № 11. С. 5–13.
39. Бондарко В.М. Сопоставление метрик, заданных на пространстве признаков изображений, с ошибками в опознании этих изображений // Сенсорные системы. 1989. Т. 3. № 1. С. 56–68.
40. Кемпбелл Ф.В., Шелепин Ю.Е. Возможности фовеолы в различении объектов // Сенсорные системы. 1990. Т. 4. № 2. С. 181–185.
41. Вахрамеева О.А., Сухинин М.В., Моисеенко Г.А. и др. Изучение порогов восприятия в зависимости от геометрии фовеа // Сенсорные системы. 2013. Т. 27. № 2. С. 122–129.
42. Chikhman V., Shelepin Y., Pronin S. Experimental study of invariant perception of wavelet images // Journal of Optical Technology. 2011. V. 78(12). P. 803–807.
43. Burt P.J. Fast filter transforms for image processing // Comput. Graph. and Image Proc. 1981. V. 16. P. 20–51.
44. Watson A.B., Ahumada A.J. A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex // IEEE Transactions on biomedical engineering. 1989. V. 36. № 1. P. 97–106.
45. Бондарко В.М., Гаузельман В.Е., Глезер В.Д. Механизмы зрительного обнаружения и опознания различны // Физиология человека. 1983. Т. 9. № 8. С. 496–498.
46. Wilson H.R., McFarlane D.K., Phillips G.C. Spatial frequency tuning of orientation selective units estimated by oblique masking // Vision Res. 1983. V. 23. № 9. P. 873–882.
47. Flom M.C. Contour interaction and the crowding effect // Problems in Optometry / Ed. by Rutstein R.P. 1991. V. 3. № 2. P. 237–257.
48. Flom M.C., Weymouth F.W., Kahneman D. Visual resolution and contour interaction // J. Opt. Soc. Am. 1963. V. 53. P. 1026–1032.
49. Bulatov A., Bertulis A. Distortions of length perception // Biol. Cybern. 1999. V. 80. P. 185–193.
50. Бондарко В.М., Бондарко Д.В., Солнушкин С.Д., Чихман В.Н. Моделирование оптических иллюзий // Оптический журнал. 2018. Т. 85. № 8. С. 5–12.
51. Bondarko V.M., Danilova M.V. Spatial interval discrimination in the presence of flanking lines // Spatial Vision. 1999. V. 12. № 2. P. 239–253.
52. Hess R.F., Badcock D.R. Metric for separation discrimination by the human visual system // J. Opt. Soc. Am. A. 1995. V. 12. P. 3–16.