

DOI: 10.17586/1023-5086-2024-91-08-99-109

УДК 612.821

Модель модулей В.Д. Глезера — возможная основа для описания зрительного восприятия

Валерия Михайловна Бондарко✉

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

vmbond@gmail.com

<https://orcid.org/0000-0001-7408-302X>

Аннотация

Предмет исследования. Исследовали механизмы зрительного восприятия. **Цель работы.** Целью настоящей работы является дальнейшее изучение возможности применения модели модулей В.Д. Глезера, описание с её помощью основных механизмов зрительного восприятия: распознавания, оценки размера и сегментации изображений. **Методы.** Использовали психофизические методы исследования и моделирование. Проведены эксперименты по распознаванию изображений, а также по оценке размеров пространственных интервалов и окружностей в иллюзиях Дельбёфа и Эббингхауза. Данные промоделированы как с использованием полных спектров изображений, так и моделью модулей, в которой изображения разлагались в конечные ряды Фурье. Для каждого изображения подбирали модуль оптимального размера, обеспечивающий сохранение наибольшего количества энергии в изображении. Рассмотрели две комбинации таких модулей. **Основные результаты.** Оказалось, что размер оптимального модуля зависит от формы изображения. Показано, что ошибки распознавания изображений и субъективные оценки их близости скоррелированы с расстояниями между изображениями, вычисленными в пространстве признаков как нормы разности спектров при совмещении у изображений центров тяжести, и частично с расстояниями, полученными в одном из вариантов модели модулей. Ранее моделью модулей были успешно аппроксимированы данные по сегментации и оценке размера изображений. Адекватность использованных моделей процессам переработки информации в зрительной системе в настоящем исследовании подтвердилась при сопоставлении полученных экспериментальных данных с проведённым анализом произведений живописи. Модель модулей может описывать механизмы оценки размера изображений и их сегментацию. Для моделирования процесса распознавания требуется усовершенствовать модель: ввести взаимодействие между модулями разных размеров. Тем самым модель модулей можно рассматривать как первое приближение к описанию зрительного восприятия. **Практическая значимость.** Модель модулей можно использовать для разработки моделей зрительного восприятия и для создания искусственных нейронных сетей, обеспечивающих сегментацию и распознавание объектов.

Ключевые слова: распознавание, оценка размера, сегментация, моделирование, пространственно-частотный анализ, модель модулей, оптические иллюзии

Благодарность: работа поддержана средствами федерального бюджета в рамках государственного задания ФГБУН Институт физиологии им. И.П. Павлова РАН (№ 1021062411653-4-3.1.8).

Ссылка для цитирования: Бондарко В.М. Модель модулей В.Д. Глезера — возможная основа для описания зрительного восприятия // Оптический журнал. 2024. Т. 91. № 8. С. 99–109. <http://doi.org/10.17586/1023-5086-2024-91-08-99-109>

Коды OCIS: 330.7326, 330.4060, 330.5510, 330.5370.

Vadim Glezer's modules model is a possible basis of vision perception description

VALERIA M. BONDARKO✉

I.P. Pavlov Institution of Physiology of the RAS, Saint-Petersburg, Russia

vmbond@gmail.com

<https://orcid.org/0000-0001-7408-302X>

Abstract

Subject of study. The mechanisms of visual perception were studied. **Aim of study** is to determine a possibility of using the modules model to describe visual perception. This model was proposed by V.D. Glezer on the basis of obtained electrophysiological data. **Method.** Psychophysical methods and modeling were used. Experiments were carried out on image recognition and size estimation of spatial intervals and circles in the Delboeuf and Ebbinghaus illusions. The experimental data were fitted by models using full image spectra and a modules model in which the images were decomposed into finite Fourier series. For each image, a module of the optimal size was selected to save the most of the energy in the image. Two combinations of such modules were considered. **Main results.** It was shown that the optimal module size depended on image. Recognition images errors and their proximity estimations correlate with the distances between images calculated in the feature space as the norm of the difference their spectra, and partially with the distances obtained in one of the variants of the modules model. The modules model adequately approximated the segmentation and size estimation data, which was confirmed by the analysis of the paintings. Thus, the module model can describe the mechanisms of size estimation and segmentation. At the same time, for recognition it is necessary to improve this model: to introduce interaction between modules of different sizes. Therefore, the module model can be considered as a first approximation to the description of visual perception. **Practical significance.** The modules model can be used to analyze images and create artificial neural networks that provide segmentation and object recognition.

Keywords: recognition, size estimation, segmentation, modeling, spatial frequency analysis, modules model, optical illusions

Acknowledgment: the study was supported by the State funding allocated to the Pavlov Institute of Physiology Russian Academy of Sciences (№ 1021062411653-4-3.1.8).

For citation: Bondarko V.M. Vadim Glezer's modules model is a possible basis of vision perception description [in Russian] // *Opticheskii Zhurnal*. 2024. V. 91. № 8. P. 99–109. <http://doi.org/10.17586/1023-5086-2024-91-08-99-109>

OCIS codes: 330.7326, 330.4060, 330.5510, 330.5370.

ВВЕДЕНИЕ

Описывать зрительное восприятие пытались многие исследователи. В настоящее время с развитием искусственных нейронных сетей эта проблема стала ещё более актуальной. Создаваемые нейронные сети направлены, в частности, на распознавание и классификацию объектов, на обнаружение искомой цели [1–4]. Многие из них используют фильтры пространственных частот (ПЧ), присутствующие в зрительной системе [5–7].

Модель модулей предложена В.Д. Глезером [8, 9] на основании полученных с коллегами электрофизиологических данных по изучению рецептивных полей (РП) нейронов зрительной коры. В исследовании [10] показано,

что часть РП при малых контрастах изображений можно считать линейными. Их реакцию на любой стимул можно представить как интеграл от весовой функции (ВФ) РП, умноженную на функцию, задающую контраст стимула. ВФ РП подобны функциям $\sin$ и $\cos$ [11] и элементам Габора [12] с количеством периодов, не превосходящем четырёх [13]. В модели осуществляется разложение изображений в конечные ряды Фурье с использованием этих функций. Основой для развития модели послужили дискретные распределения размеров РП и оптимальных ПЧ, на которые настроены нейроны коры [13]. Модель подобна модели пирамид с дискретным и конечным набором ПЧ фильтров [14].

В предыдущих наших исследованиях [15, 16] была предпринята попытка моделью модулей аппроксимировать экспериментальные данные по оценке размера и сегментации изображений. Оказалось, что изображения разной формы кажутся равными по размеру, если модули, оптимально их описывающие (сохраняющие максимум энергии у изображений при ограниченном числе фильтров), имеют одинаковый размер. Сегментация происходит успешно при окружении изображений однородным полем, не меньшем по размеру оптимального модуля. Моделью аппроксимировали данные по оценке размеров в иллюзиях Эббингхауза и Оппель–Кундта и по выравниванию размеров изображений прямых и косых крестов, которые, кроме того, использовали в экспериментах по сегментации. Выводы работ дополнили рассмотрением работ живописи и их композиции, которые проанализировали с точки зрения относительных размеров их отдельных частей.

Целью настоящей работы является дальнейшее изучение возможности применения модели модулей, описание с её помощью основных механизмов зрительного восприятия: распознавания, оценки размера и сегментации изображений.

В исследовании будут представлены новые экспериментальные данные по оценке этих функций и проведено их сопоставление с вычислениями модели. Успешная аппроксимация полученных в экспериментах результатов позволит рассматривать модель модулей как основу описания процесса зрительного восприятия.

Моделирование распознавания зрительных изображений ПЧ фильтрами было предпринято вследствие предложенной Кемпбеллом, Робсоном и Блэйкмором [17, 18] гипотезы о ПЧ анализе изображений, осуществляемым зрительной системой. Использовали известные спектральные модели. Это модель согласованной фильтрации и модель, в которой разницу между спектрами оценивали при совмещении у изображений центров тяжести. Высокие значения корреляции между ответами наблюдателей и расстояниями между изображениями, вычисленными по этим моделям [19–21], свидетельствуют об адекватности моделей. В настоящем исследовании будут сравниваться результаты, полученные при использовании

этих спектральных моделей и модели модулей. Ранее такое сопоставление не проводилось.

В модели модулей используются фильтры разного размера и настройки, как и в свёрточных нейронных сетях [22–26]. Полученные в настоящем исследовании результаты могут способствовать дальнейшему усовершенствованию разработки нейронных сетей и моделей зрительного восприятия.

МЕТОДИКА

Стимулы. В экспериментах по распознаванию стимулами служили шесть изображений (рис. 1), вписанных в окружности одного диаметра. При кратковременном предъявлении их размер был равен 1,1 угл. град. Время экспозиции подбирали так, чтобы вероятность правильного распознавания у каждого наблюдателя была около 0,5. Те же изображения использовали для оценки их схожести с неограниченным временем предъявления и размером 2–3 угл. град.

При оценке размера в качестве референтных использовали изолированные стимулы или стимулы с симметричным окружением (рис. 2). В одних экспериментах сравнивали ширину прямоугольных участков плоскости, ограниченных двумя вертикальными линиями, так называемых пространственных интервалов (*spatial intervals*), в других — величину окружностей. Интервалы окружали с двух сторон двумя вертикальными линиями (рис. 2д). К окружностям с четырёх сторон добавляли либо четыре одинаковые окружности, либо линии, равные по длине диаметру окружности. Меняли расстояние между центральным изображением и окружением. Диаметр дополнительных окружностей был в два раза больше или меньше центральной (иллюзия Эббингхауза — рис. 2а, 2б). Кроме того, дополнительным изображением являлась

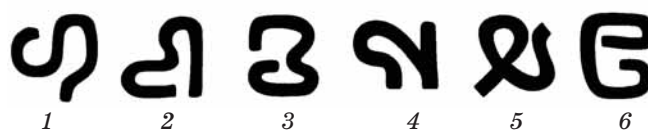


Рис. 1. Алфавит изображений. Цифры использовали для ответа

Fig. 1. Image alphabet. The numbers were used for the answer

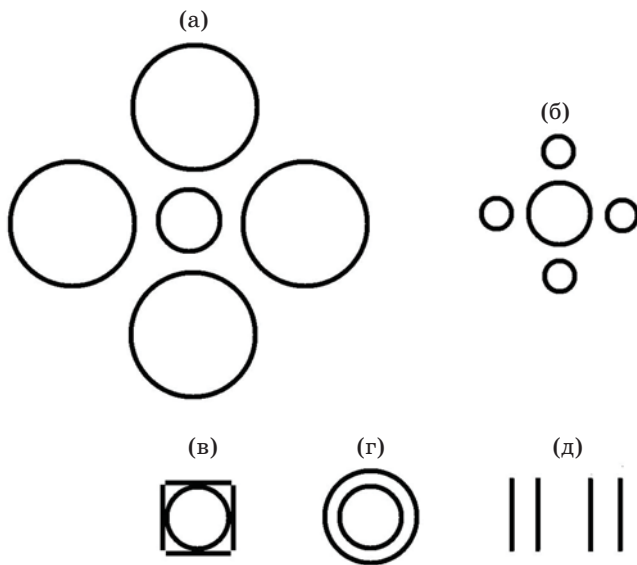


Рис. 2. Референтные стимулы с окружением. (а)–(г) — центральные окружности одинакового размера в иллюзиях Эббингхауса (а, б) и Дельбёфа (г). (д) — пространственный интервал с боковыми полосами

Fig. 2. Referent stimuli with surroundings. (а)–(г) — central circles of the same size in the Ebbinghaus (а, б) and Delboeuf (г) illusions. (д) — spatial interval with side bars

окружность варьируемого большего диаметра (иллюзия Дельбёфа — рис. 2г). Тестовыми стимулами были пространственные интервалы или окружности, у которых менялся размер.

Процедура. Стимулы предъявляли бинокулярно. Применялась методика вынужденного выбора. Ответ «не знаю» разрешён не был. В экспериментах по распознаванию наблюдателей предварительно знакомили с изображениями. Их предъявляли по одному в случайном порядке. Регистрировали ответы, по которым строили матрицы ответов A_{ij} : при предъявлении i -го стимула количество ответов «вижу j -ый стимул». Анализировали суммарные данные каждого наблюдателя по 10 опытам (450 предъявлений отдельного стимула).

Карточки с теми же изображениями раздавали испытуемым для проведения субъективной оценки их схожести по пятибалльной шкале. Самым похожим изображениям ставилась оценка «1», самым непохожим — «5». Матрицы субъективных оценок суммировали по данным всех наблюдателей.

При оценке размеров наблюдатели отвечали на вопрос, больше или меньше референт-

ный стимул тестового. Референтные и тестовые стимулы предъявляли одновременно справа или слева на расстоянии 3 угл. град от центра экрана. Размер центральных изображений у референтных стимулов 21 угл. мин, ширина линий 1,5 угл. мин. Расстояния до дополнительных изображений меняли случайным образом. При фиксированном размере референтного стимула меняли размеры тестовых изображений. Шаг изменения подбирался индивидуально для каждого наблюдателя так, чтобы вероятность ответов варьировалась в пределах 0,1–0,9 ед. Использовали 7 шагов. Каждая пара изображений была предъявлена не менее 50 раз.

Аппаратура. При оценке размера стимулы предъявляли на мониторе LG Flatron 795FT Plus с размером диагонали 17" при разрешении экрана 1024×768 пикселей с частотой вертикальной развёртки 100 Гц на расстоянии 3,5 м. Яркость экрана — 90 кд/м², изображений — 5 кд/м², время предъявления — 1 с. При распознавании использовали монитор Diamond Plus 230SB с размером ЭЛТ 22" с разрешением 1024×768 пикселей и яркостью экрана 40 кд/м². Время предъявления — 30–50 мс, контраст изображений — 0,03, расстояние до экрана — 115 см. Для калибровки монитора применяли фотометр фирмы Cambridge Research Systems. Программное обеспечение реализовано в среде Delphi 6.

Моделирование. Предъявляемые в экспериментах изображения были оцифрованы: представлены в виде двумерных матриц. Во вход модели модулей подавались оцифрованные изображения, которые разлагались в двумерные конечные ряды Фурье с помощью функций $\sin$ и $\cos$ на квадратных участках, больших или равных по размеру изображениям. Размеры участков разложения варьировали. Для каждого изображения подбирали размер оптимального модуля таким образом, чтобы сохранить наибольшее количество энергии у изображения. Рассмотрены модули с 3, 4 или 5 гармониками. За начало отсчёта выбиралась точка, в которой коэффициенты при нечётной функции на первой гармонике были равны нулю. Выходом модели служили коэффициенты разложения изображений в ряды Фурье.

Данные по распознаванию сопоставлялись с вычисленными расстояниями между изо-

бражениями как нормами разности спектров при максимальном их совпадении (модель согласованной фильтрации — МСФ) или при совмещении у изображений центров тяжести (МЦТ), а также в модели модулей (ММ) как нормами разности коэффициентов разложения.

Обработка данных. Сравнение данных испытуемых по распознаванию и схожести изображений с вычисленными в моделях расстояниями проведено посредством корреляционного анализа. Пробит анализ использовали при оценке размера: психометрические функции для ответов «больше» (суммарные ответы по всем экспериментам при каждом размере тестового изображения) методом наименьших квадратов приближали к функциям нормального распределения. Средние значения нормальных распределений соответствуют тем размерам, при которых наблюдатели считают их равными тестовым.

Наблюдатели. В экспериментах по распознаванию приняло участие трое наблюдателей в возрасте 27–38 лет с нормальным зрением. Шкалирование осуществили 52 наблюдателя, оценку размера в длительных экспериментах — 2 наблюдателя в возрасте 35 и 42 лет и 7 наблюдателей с меньшим набором референтных стимулов. Исследование было одобрено Этическим комитетом СПбГУ.

РЕЗУЛЬТАТЫ

Распознавание изображений. Полученные в экспериментах матрицы с распределением ответов и оценкой их сходства были сопоставлены с расстояниями между изображениями, вычисленными по четырём моделям. Оказалось, что каждое изображение лучше описывается модулем своего размера. Размеры оптимальных модулей варьировались от 1,2 до 2,1 относительно размеров изображений. Поэтому были взяты две разновидности модели модулей. В модели «среднего» модуля (МСрМ) размеры оптимальных модулей усреднили. Средний размер модуля оказался равен 1,5 размерам изображений. Для такого модуля для каждого изображения были рассчитаны коэффициенты двумерного ряда Фурье. Расстояния вычислялись как корни квадратов из сумм квадратов разностей коэффициентов. В модели «своего» модуля (МСвМ) сравнивались коэффициенты разложения, полу-

ченные для оптимальных размеров модулей каждого изображения.

В табл. 1 в строках сначала приведены строки из матрицы ответов одного из наблюдателей (при предъявлении стимула, указанного в строке, количество ответов на стимул, указанный в столбце — Отв), затем субъективные оценки близости (СОБ) изображений, далее нормированные расстояния, вычисленные по МЦТ и МСрМ с четырьмя гармониками. Расстояния в МСвМ и МСФ не показаны из-за меньшего соответствия их экспериментальным данным (табл. 2).

Ответы наблюдателя, оценки сходства изображений и расстояния между ними взаимосвязаны. Например, наблюдатель чаще называет первое изображение четвёртым или пятым; более похожим на первое изображение наблюдатели также считают четвёртый стимул (меньшее количество баллов); между первым и четвёртым стимулом и минимальные расстояния. Аналогичную картину можно увидеть и при сравнении третьего и шестого изображений и т.д. Для анализа результатов посчитаны корреляции между полными матрицами и между матрицами без диагоналей, поскольку оказалось, что диагонали дают больший вклад из-за существенной разницы с остальными элементами. Данные по распознаванию в матрице без диагоналей соответствуют ошибочным ответам. Корреляции для полных матриц и матриц без диагоналей приведены в табл. 2 в верхних и нижних строках каждой ячейки.

Полные матрицы с количеством правильных ответов дают высокие корреляции с любыми расстояниями, что делает невозможным по ним определение преимущества какой-либо из моделей. Корреляции между ошибками наблюдателей и расстояниями отрицательные: чем меньше расстояние в пространстве признаков, тем больше ошибок. Двумя звёздочками отмечены корреляции, достоверные на уровне значимости 1%, одной — на уровне 5%. Скоррелированы со всеми экспериментальными данными расстояния, вычисленные в спектральной модели с совмещением у изображений центров тяжести. В то же время у двух наблюдателей достоверный результат дают модели согласованной фильтрации и среднего модуля. Такое расхождение нуждается в пояснении. Процесс распознавания индивидуален. Большая корреляция в МСФ

Таблица 1. Сопоставление экспериментальных данных с результатами моделирования**Table 1.** Comparison of experimental data with modeling results

Номера стимулов	Количество ответов (Отв) наблюдателя (вижу j -ый стимул при предъявлении i -го стимула), субъективные оценки близости изображений (СОБ), расстояния в моделях МЦТ и МСрМ					
	1	2	3	4	5	6
Отв	185	46	37	65	86	31
СОБ	0	126	171	111	207	223
1 МЦТ	0	0,95	1	0,77	0,86	0,94
МСрМ	0	0,89	1	0,79	0,83	0,86
Отв	46	291	38	18	33	24
СОБ	123	0	122	134	239	179
2 МЦТ	0,95	0	0,96	0,9	0,9	0,95
МСрМ	0,89	0	0,9	0,89	0,88	0,87
Отв	31	38	225	19	44	93
СОБ	163	115	0	207	238	98
3 МЦТ	1	0,96	0	0,95	0,99	0,8
МСрМ	1	0,9	0	0,96	0,86	0,83
Отв	81	34	19	221	78	17
СОБ	108	136	209	0	162	231
4 МЦТ	0,77	0,9	0,95	0	0,88	0,97
МСрМ	0,79	0,89	0,96	0	0,9	0,89
Отв	55	61	49	62	172	51
СОБ	128	187	212	138	0	238
5 МЦТ	0,86	0,9	0,99	0,88	0	0,98
МСрМ	0,83	0,88	0,86	0,9	0	0,85
Отв	32	23	64	15	49	267
СОБ	179	152	90	213	237	0
6 МЦТ	0,94	0,95	0,8	0,97	0,98	0
МСрМ	0,86	0,87	0,83	0,89	0,85	0

Таблица 2. Корреляции между ответами отдельных наблюдателей и оценками близости изображений с расстояниями, вычисленными по разным моделям**Table 2.** Correlations between individual observers' responses and proximity ratings of images with distances calculated in different models

Модель	Наблюдатель S1	Наблюдатель S2	Наблюдатель S3	Субъективное шкалирование
1. МЦТ	−0,98 −0,65**	−0,96 −0,37*	−0,95 −0,68**	0,86 0,58**
2. МСФ	−0,98 −0,37*	−0,96 −0,46**	−0,93 −0,22	0,83 0,2
3. МСвМ	−0,79 −0,22	−0,78 −0,21	−0,78 −0,26	0,74 0,27
4. МСрМ	−0,98 −0,37*	−0,96 −0,2	−0,95 −0,56**	0,83 0,18

по сравнению с МЦТ у наблюдателя S2 может быть обусловлена тем, что этот наблюдатель был хорошо знаком с алфавитом. У наблюдателей S1 и S3 высокие корреляции в МЦТ, но и достоверные в МСрМ. Сравним эти модели. В табл. 3 показаны корреляции между расстояниями, вычисленными в моделях.

Скоррелированы расстояния в МЦТ и МСрМ, а также в МСФ и МСвМ. Таким образом, достоверные корреляции у наблюдателей S1 и S3 в МЦТ и МСрМ не удивительны. Также кажется оправданным и то, что субъективные оценки не скоррелированы ни с одной из моделей модулей. Модель модулей не обрабатывает высокие частоты в изображениях. При кратковременном предъявлении высо-

кие ПЧ не анализируются зрительной системой [8, 9, 26], поэтому у наблюдателей S1 и S3 данные согласованы и с МЦТ, и с МСрМ. Субъективное шкалирование проведено без ограничения времени, здесь высокие частоты вносят свой вклад в восприятие изображений.

Оценка размера. Отклонения в оценке размера изображений (иллюзии) в зависимости от расстояний между референтными стимулами и окружением у двух наблюдателей показаны на рис. 3а и 3б. Они вычислены как разницы между средними значения нормальных распределений, аппроксимировавших ответы наблюдателей, и физическими размерами стимулов. Прослеживаются закономерности. Наибольшее преувеличение размера выявлено в иллюзии Дельбёфа (кривые 1), затем в оценке ширины пространственных интервалов (кривые 2), далее в присутствии дополнительных линий, окружностей меньшего и большего диаметра (кривые 3, 4 и 5 — иллюзия Эббингхауза). Но в то же время есть и различия в данных наблюдателей. На рис. 3б величины иллюзий меньше, и они становятся противоположными по знаку при меньших расстояниях.

Различная зависимость от расстояния привела к необходимости провести дополнительные исследования на группе наблюдателей с меньшим числом стимулов и расстояний в ил-

Таблица 3. Корреляции между расстояниями, вычисленными в разных моделях
Table 3. Correlations between distances calculated in different models

	МЦТ	МСрМ	МСФ	МСвМ
МЦТ	1			
МСрМ	0,64**	1		
МСФ	0,35	0,18	1	
МСвМ	0,20	−0,12	0,54**	1

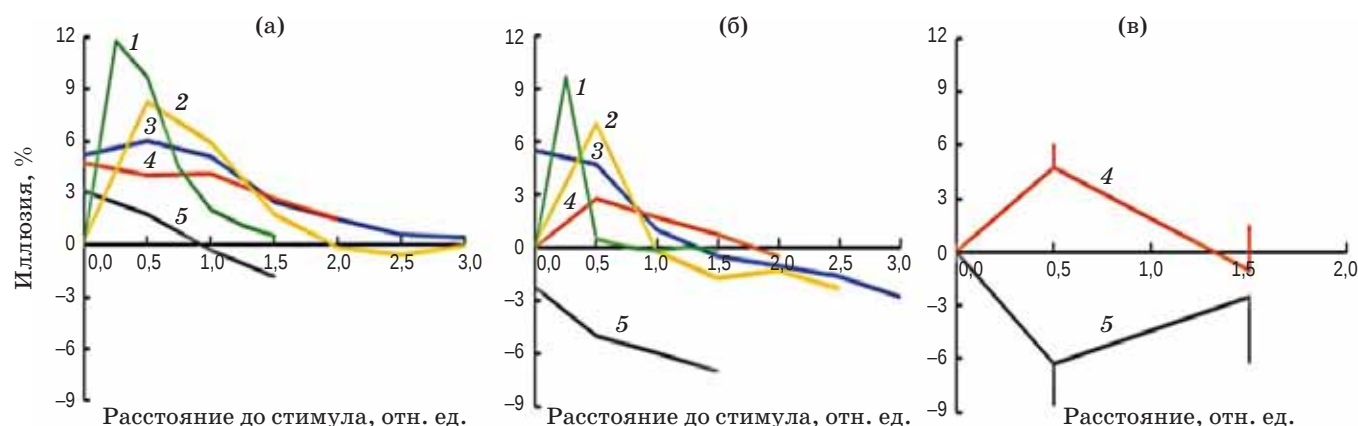


Рис. 3. Оценка размера изображения двумя наблюдателями (а, б) в иллюзии Дельбёфа (кривые 1) и Эббингхауза при боковых линиях и окружностях вдвое меньшего и большего диаметров (кривые 3, 4, 5), пространственного интервала (кривые 2). (в) — усреднённые оценки в иллюзии Эббингхауза у группы наблюдателей при окружностях, вдвое меньшего и большего диаметров (кривые 4, 5). Вертикальные отрезки — стандартные ошибки средних

Fig. 3. Estimation of size by two observers (a, б) in the Delboeuf illusion (curves 1) and Ebbinghaus illusion with lateral lines and circles of half the smaller and larger diameters (curves 3, 4, 5), spatial interval (curves 2). (v) — averaged estimates in the Ebbinghaus illusion for a group of observers with circles half the size and twice the diameter (curves 4, 5). Vertical bars are standard errors of the means

люзии Эббингхауза с окружностями вдвое меньшего и большего размера (рис. 3в). Данные свидетельствуют о вариабельности оценок у наблюдателей, как ранее было показано у наблюдателей разного возраста [28].

ОБСУЖДЕНИЕ

Распознавание изображений и модель модулей. Ошибки распознавания и оценки сходства изображений оказались скоррелированными с расстояниями в пространстве признаков, вычисленными по спектральной модели. К аналогичным выводам пришли и другие исследователи [19–21], использовавшие в экспериментах различные изображения. Модель среднего модуля с размером в 1,5 раза превосходящем размер изображений, показала менее приемлемый результат, видимо, из-за того, что в ней отсутствует анализ изображений по высоким частотам. Попытки определить оптимальные модули меньшего размера, чем изображения, также были предприняты в настоящей работе. Оказалось, что оптимальными в этом случае являются модули с размером, примерно в два раза большим ширины линий в стимулах. Но каким образом нужно объединять модули разного размера, неизвестно. Усовершенствованную на основании экспериментально полученного распределения размеров с разницей в половину октавы 1,0; 1,41; 2,0; 2,83 и т. д. модель модулей [12] использовали только для инвариантного к размеру описания изображений. Алгоритм взаимодействия модулей в ней не указан. Тем не менее полученная согласованность модели модулей и спектральной модели позволяет предположить, что её можно взять за основу описания процесса распознавания, т.к. она основана на экспериментальных данных.

Оценка размера и сегментация. В экспериментах выявлены закономерности в оценке размеров: при близком расположении окружения к стимулу чаще наблюдается преувеличение размера объекта. Поскольку наибольшая переоценка размера возникает в иллюзии Дельбёфа, можно полагать, что происходит ассимиляция стимулов с окружением (в иллюзии Дельбёфа окружение располагается ближе всего к стимулу). Оценки размеров в других экспериментах этому не противоречат. По мере удаления дополнительных изображений иллюзии исче-

зают или меняют знак, т.е. эти изображения при больших расстояниях оказывают тормозное воздействие на восприятие центрального стимула. Таким взаимодействием можно объяснить и преуменьшение размера в иллюзии Эббингхауза для окружностей вдвое большего диаметра (рис. 3, кривые 5). Ранее были получены схожие оценки в этой иллюзии [16] при использовании дополнительных окружностей того же диаметра, вдвое меньшего диаметра или линий, что противоречит гипотезе о возникновении иллюзии вследствие контрастности размеров изображений и предполагает наличие механизма взаимодействия контуров [28, 29].

В предыдущей работе [16] показано, что оценка размеров связана с сегментацией изображений. Сегментация осуществляется, когда дополнительные изображения находятся на определённом расстоянии от тестируемого стимула, то есть стимул должен быть окружён однородным полем без дистракторов. Результаты этих экспериментов аппроксимированы моделью модулей. Отношения размеров оптимальных модулей и минимальных областей, при которых происходит сегментация, к размерам изображений практически совпали. Относительные размеры оптимальных модулей оказались близкими к 1,5 ед., что позволило сделать вывод о том, что при сегментации окружение выходит за пределы оптимального модуля и описывается уже модулем следующего размера. Кроме того, изображения разной формы казались равными по размеру, когда были равны размеры их оптимальных модулей.

В настоящем исследовании эксперименты по сегментации не проводились. Минимальный размер области, где нет влияния окружения, скорее всего, связан с сегментацией, но его трудно определить из-за большого различия в оценках наблюдателей. Так, у одного наблюдателя иллюзия Дельбёфа исчезает при расстоянии до дополнительного изображения, равного половине диаметра, то есть в присутствии окружности вдвое большей исходной (рис. 3б, кривая 1). У другого наблюдателя влияние дополнительного изображения проявляется на всех расстояниях. Усреднённые данные семи наблюдателей по изучению иллюзии Эббингхауза также не дают ответа на вопрос о зоне взаимодействия из-за большого

разброса данных (рис. 3в). Поэтому представляет интерес сопоставление полученных данных с произведениями живописи, которые учитывают, как правило, основные закономерности зрительного восприятия.

Сопоставление с произведениями живописи. В предыдущих наших исследованиях [15, 16] размеры изображений и окружающих их полей в картинах Казимира Малевича «Чёрный квадрат» и «Чёрный круг» сравнивались с результатами вычислений в модели модулей и данными по сегментации и оценке размера. Соотношения размеров в картинах оказались согласованными с моделью модулей. В картине из Третьяковской галереи внешний размер рамки в 1,4 раза превосходит размер квадрата. Согласно модели модулей [12], такое соотношение обеспечивает попадание рамки в следующий по размеру модуль. Рамка в «Чёрном круге» располагается несимметрично по отношению к кругу, однако при таком расположении существенно преувеличивает его размер (рис. 3, кривая 3). То есть, в картинах с простыми формами, которые мы анализировали из-за удобства измерений изображений, К. Малевич интуитивно использовал закономерности зрительного восприятия для усиления впечатления.

Произведения живописи обычно более сложные и многоплановые. В нашей работе [16] также была рассмотрена общепринятая композиция иконы Андрея Рублева «Троица». Она представляет собой ряд концентрических окружностей с соотношением диаметров, близким к золотому сечению (1,62) и к разнице в размерах модулей. В предыдущих исследованиях оценку размера концентрических окружностей (иллюзию Дельбёфа) не изучали. Между тем в иконописи принято изображать нимбы с диаметром, в полтора-два раза большим размеров ликов святых, что согласуется с полученными в настоящем исследовании данными по оценке размера в иллюзии Дельбёфа. При таком соотношении наблюдается максимальное преувеличение размера. Можно привести множество примеров использования кругов и овалов или их частей для подчеркивания (сегментации) и рассматривания отдельных планов картины и изображений как в иконописи, так и в светской живописи. Соотношения их размеров согласованы с результатами настоящего исследования и с дискретизацией размеров модулей. Примером могут служить зна-

менитые фрески Дионисия в Ферапонтовом монастыре, Рафаэля в Сикстинской капелле («Афинская школа») и другие. Поэтому можно использовать модель модулей для описания механизмов оценки размера и сегментации. Необходимо подчеркнуть, что художники всегда уделяли большое внимание и к обрамлению картин, что также способствует выделению объектов из фона.

ЗАКЛЮЧЕНИЕ

Впервые показано, что описание изображений с наименьшей потерей информации конечным набором ПЧ фильтров, реализованное в модели модулей, близко к спектральному с полным набором частот. Размеры оптимальных модулей превосходит объекты в среднем в 1,5 раза. Для описания распознавания необходимо вводить взаимодействие между модулями разных размеров в то время, как сегментацию и оценку размера могут осуществлять оптимальные модули без учёта модулей других размеров. Поэтому модель модулей можно рассматривать как первое приближение к описанию процесса зрительного восприятия.

Проиллюстрирована согласованность между экспериментальными данными и композициями произведений живописи. Этот факт, а также выявленная корреляция между результатами психофизических экспериментов по зрительному восприятию и моделью модулей подтверждают адекватность этой модели реальным процессам обработки изображений, происходящим в зрительной системе. Следует отметить, что модель модулей разработана на основе полученных электрофизиологических данных, что является убедительным дополнительным аргументом в её пользу. Близость модели к иерархически организованным моделям пирамид [14], служащим для передачи и сжатия информации, предоставляет возможность для дальнейшего развития модели, использовании её в алгоритмах искусственных свёрточных нейронных сетей, осуществляющих анализ изображений, их сегментацию и распознавание. Возможно, что использование нейронных сетей с ядрами, отличающимися по размеру на полоктавы и имеющими в каждом слое больший набор весовых функций, окажется более эффективным для успешного применения их на практике.

СПИСОК ИСТОЧНИКОВ

1. Явна Д.В., Бабенко В.В., Горбенкова О.А. и др. Категоризация объектов и сцен нейронной сетью, входы которой предварительно обучены декодированию пространственных неоднородностей текстуры // *Оптический журнал*. 2023. Т. 90. № 1. С. 37–48. <https://doi.org/10.17586/1023-5086-2023-90-01-37-48>
2. Ермаченкова М.К., Малашин Р.О., Бойко А.А. Обучение нейронных сетей для классификации тепловизионных изображений на основе изображений видимого спектра // *Оптический журнал*. 2023. Т. 90. № 10. С. 48–66. <https://doi.org/10.17586/1023-5086-2023-90-10-48-66>
3. Цыцулин А.К., Бобровский А.И., Морозов А.В. и др. 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. <https://doi.org/10.17586/1023-5086-2019-86-10-30-38>
4. Жукова О.В., Малахова Е.Ю., Шелепин Ю.Е. Джоконда и неопределённость распознавания улыбки человеком и искусственной нейронной сетью // *Оптический журнал*. 2019. Т. 86. № 11. С. 40–50. <https://doi.org/10.17586/1023-5086-2019-86-11-40-50>
5. Луцив В.Р. Свёрточные искусственные нейронные сети глубокого обучения // *Оптический журнал*. 2015. Т. 82. № 8. С. 11–23.
6. Малахова Е.Ю. Пространство описания зрительной сцены в искусственных и биологических нейронных сетях // *Оптический журнал*. 2020. Т. 87. № 10. С. 50–58. <https://doi.org/10.17586/1023-5086-2020-87-10-50-58>
7. Малахова Е.Ю. Представление категорий посредством прототипов согласованной активности нейронов в свёрточных нейронных сетях // *Оптический журнал*. 2021. Т. 88. № 12. С. 36–41. <https://doi.org/10.17586/1023-5086-2021-88-12-36-41>
8. Глезер В.Д. Зрение и мышление. Л.: Наука, 1985. 300 с.
9. Glezer V.D. Vision and mind: Modeling mental functions. NJ.: Lawrence Erlbaum Ass., 1995. 274 p.
10. Glezer V.D., Tcherbach T.A., Gauzelman V.E., Bondarko V.M. Linear and non-linear properties of simple and complex receptive fields in area 17 of the cat visual cortex: A model of the fields // *Biol. Cybern.* 1980. V. 37. P. 195–208.
11. Вол И.А. Пространственно-частотная модель гипертропии зрительной системы // *Сенсорные системы*. 1988. Т. 2. № 2. С. 133–138.
12. 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. <https://doi.org/10.1364/AO.33.005256>
13. 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. <https://doi.org/10.1017/s0952523800005526>
14. Burt P.J. Fast filter transforms for image processing // *Comput. Graph. and Image Proc.* 1981. V. 16. P. 20–51.
15. Бондарко В.М., Данилова М.В., Чихман В.Н. Сегментация зрительных изображений: экспериментальные данные и моделирование // *Оптический журнал*. 2021. Т. 88. № 12. С. 7–17. <http://doi.org/10.17586/1023-5086-2021-88-12-7-17>
16. Бондарко В.М., Данилова М.В. Связь размера локального окна в модели модулей с оценкой размера зрительных изображений и их сегментацией // *Оп-*

REFERENCES

1. Yavna D.V., Babenko V.V., Gorbenkova O.A. et al. Classification of objects and scenes by a neural network with pretrained input modules to decode spatial texture inhomogeneities // *Journal of Optical Technology*. 2023. V. 90(1). P. 20–25. <https://doi.org/10.1364/JOT.90.000020>
2. Ermachenkova M.K., Malashin R.O., Boyko A.A. Training of neural networks for classification of thermal imaging based on images of the visible spectrum [in Russian] // *Journal of Optical Technology*. 2023. V. 90(10). P. 590–600. <https://doi.org/10.1364/JOT.90.000590>
3. Tsytulin A.K., Bobrovskiy A.I., Morozov A.V. et al. Using convolutional neural networks to automatically select small artificial space objects on optical images of a starry sky // *Journal of Optical Technology*. 2019. V. 86(10). P. 627–633. <https://doi.org/10.1364/JOT.86.000627>
4. Zhukova O.V., Malakhova E.Yu., Shelepin Yu.E. La Gioconda and the indeterminacy of smile recognition by a person and by an artificial neural network // *Journal of Optical Technology*. 2019. V. 86(11). P. 706–715. <https://doi.org/10.1364/JOT.86.000706>
5. Lutsiv V.R. Convolutional artificial neural networks of deep learning [in Russian] // *Journal of Optical Technology*. 2015. T. 82. № 8. P. 499–508. <https://doi.org/10.1364/JOT.82.000499>
6. Malakhova E.Yu. Information representation space in artificial and biological neural networks // *Journal of Optical Technology*. 2020. V. 87(10). P. 598–603. <https://doi.org/10.1364/JOT.87.000598>
7. Malakhova E.Yu. Representation of categories through prototypes formed based on coordinated activity of units in convolutional neural networks // *Journal of Optical Technology*. 2021. V. 88. № 12. P. 706–709. <https://doi.org/10.1364/JOT.88.000706>
8. Glezer V.D. Vision and thinking [in Russian]. L.: Nauka, 1985. 300 p.
9. Glezer V.D. Vision and mind: Modeling mental functions. NJ.: Lawrence Erlbaum Ass., 1995. 274 p.
10. Glezer V.D., Tcherbach T.A., Gauzelman V.E., Bondarko V.M. Linear and non-linear properties of simple and complex receptive fields in area 17 of the cat visual cortex: A model of the fields // *Biol. Cybern.* 1980. V. 37. P. 195–208.
11. Vol I.A. Spatial-frequency model of hyperacuity of the visual system [in Russian] // *Sensory systems*. 1988. V. 2. № 2. P. 133–138.
12. 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.
13. 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.
14. Burt P.J. Fast filter transforms for image processing // *Comput. Graph. and Image Proc.* 1981. V. 16. P. 20–51.
15. Bondarko V.M., Danilova M.V., Chikhman V.N. Segmentation of visual images: experimental data and modelling // *J. Opt. Technol.* 2021. V. 88(12). P. 692–699. <https://doi.org/10.1364/JOT.88.000692>
16. Bondarko V.M., Danilova M.V. Relation of local window size in a model of modules with estimation of the size of visual images and their segmentation // *J. Opt. Technol.* 2022. V. 89(8). P. 461–468. <https://doi.org/10.1364/JOT.89.000461>

- тический журнал. 2022. Т. 89. № 8. С. 43–53. <https://doi.org/10.17586/1023-5086-2022-89-08-43-53>
17. Blakemore G., Campbell F.W. On the existence in human visual system of neurones selectively sensitive to the orientation and size of retinal image // *J. Physiology*. 1969. V. 203. № 1. P. 237–260.
18. Campbell F.W., Robson J.G. Application of Fourier analyses to the visibility of gratings // *J. Physiol*. 1968. V. 197. P. 551–566.
19. Вол И.А., Павловская М.Б. Корреляция между близостью фурье-спектров изображений и ошибками распознавания // *Физиол. чел.* 1986. Т. 12. № 3. С. 400–406.
20. Gervais M. J., Harvey L.O., Roberts Y.O. Identification confusions among letters of the alphabet // *J. Exp. Psychol. Hum. Percept.* 1984. V. 10. № 5. P. 655–666.
21. Vol I.A., Pavlovskaya M.B., Bondarko V.M. Similarity between Fourier transform of images predicts their experimental confusion // *Perception and Psychophysics*. 1990. V. 47. № 1. P. 12–21.
22. Patil A., Rane M. Convolutional neural networks: an overview and its applications in pattern recognition // *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020*. 2021. V. 1. P. 21–30. https://doi.org/10.1007/978-981-15-7078-0_3
23. Rawat W., Wang Z. Deep convolutional neural networks for image classification: A comprehensive review // *Neural computation*. 2017. V. 29. № 9. P. 2352–2449. https://doi.org/10.1162/neco_a_00990
24. Sharma N., Jain V., Mishra A. An analysis of convolutional neural networks for image classification // *Procedia computer science*. 2018. V. 132. P. 377–384. <https://doi.org/10.1016/j.procs.2018.05.198>
25. Титаренко М.А., Малашин Р.О. Исследование способностей нейронных сетей к извлечению и использованию семантической информации при обучении восстановлению зашумлённых изображений // *Оптический журнал*. 2022. Т. 89. № 2. С. 25–35. <https://doi.org/10.17586/1023-5086-2022-89-02-25-35>
26. Альес М.Ю., Антонов Е.А., Калугин А.И., Зарипов М.Р. Применение искусственных нейронных сетей для анализа мультиспектральных изображений // *Оптический журнал*. 2021. Т. 88. № 8. С. 48–53. <https://doi.org/10.17586/1023-5086-2021-88-08-48-53>
27. Mihaylova M., Stomonyakov V., Vassilev A. Peripheral and central delay in processing high spatial frequencies: reaction time and VEP latency studies // *Vision research*. 1999. V. 39. № 4. P. 699–705. [https://doi.org/10.1016/S0042-6989\(98\)00165-5](https://doi.org/10.1016/S0042-6989(98)00165-5)
28. Бондарко В.М., Семенов Л.А. Оценка размера в иллюзии Эббингхауса у взрослых и детей различного возраста // *Физиология человека*. 2004. Т. 30. № 1. С. 31–37.
29. Weintraub D.J., Schneek M.K. Fragments of Delboeuf and Ebbinghaus illusions: Contour context explorations of misjudged circle size // *Perception and Psychophysics*. 1986. V. 40 (3). P. 147–158.
17. Blakemore G., Campbell F.W. On the existence in human visual system of neurones selectively sensitive to the orientation and size of retinal image // *J. Physiology*. 1969. V. 203. № 1. P. 237–260.
18. Campbell F.W., Robson J.G. Application of Fourier analyses to the visibility of gratings // *J. Physiol*. 1968. V. 197. P. 551–566.
19. Vol I.A., Pavlovskaya M.B. Correlation between proximity of Fourier spectra of images and errors in their identification // *Human Physiology*. 1986. V. 12. P. 400–406.
20. Gervais M. J., Harvey L.O., Roberts Y.O. Identification confusions among letters of the alphabet // *J. Exp. Psychol. Hum. Percept.* 1984. V. 10. № 5. P. 655–666.
21. Vol I.A., Pavlovskaya M.B., Bondarko V.M. Similarity between Fourier transform of images predicts their experimental confusion // *Perception and Psychophysics*. 1990. V. 47. № 1. P. 12–21.
22. Patil A., Rane M. Convolutional neural networks: an overview and its applications in pattern recognition // *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020*. 2021. V. 1. P. 21–30. https://doi.org/10.1007/978-981-15-7078-0_3
23. Rawat W., Wang Z. Deep convolutional neural networks for image classification: A comprehensive review // *Neural computation*. 2017. V. 29. № 9. P. 2352–2449. https://doi.org/10.1162/neco_a_00990
24. Sharma N., Jain V., Mishra A. An analysis of convolutional neural networks for image classification // *Procedia computer science*. 2018. V. 132. P. 377–384. <https://doi.org/10.1016/j.procs.2018.05.198>
25. Titarenko M.A., Malashin R.O. Study of the ability of neural networks to extract and use semantic information when they are trained to reconstruct noisy images // *Journal of Optical Technology*. 2022. V. 89(2). P. 81–88. <https://doi.org/10.1364/JOT.89.000081>
26. Alies M.Yu., Antonov E.A., Kalugin A.I., Zaripov M.R. Application of artificial neural networks for the analysis of multispectral images // *Journal of Optical Technology*. 2021. V. 88(8). P. 441–444. <https://doi.org/10.1364/JOT.88.000441>
27. Mihaylova M., Stomonyakov V., Vassilev A. Peripheral and central delay in processing high spatial frequencies: reaction time and VEP latency studies // *Vision research*. 1999. V. 39. № 4. P. 699–705. [https://doi.org/10.1016/S0042-6989\(98\)00165-5](https://doi.org/10.1016/S0042-6989(98)00165-5)
28. Bondarko V.M., Semenov L.A. Size estimation in the Ebbinghaus illusion in adults and children of different ages // *Human Physiology*. 2004. V. 30. P. 24–30.
29. Weintraub D.J., Schneek M.K. Fragments of Delboeuf and Ebbinghaus illusions: Contour context explorations of misjudged circle size // *Perception and Psychophysics*. 1986. V. 40 (3). P. 147–158.

АВТОРЫ

Валерия Михайловна Бондарко — доктор биол. наук, ведущий научный сотрудник, Институт физиологии им. И.П. Павлова РАН, 199034, Санкт-Петербург, Россия; Scopus ID: 7004163777, <https://orcid.org/0000-0001-7408-302X>, vmbond@gmail.com

AUTHORS

Valeria M. Bondarko — Dr.Sc. in Biology, Leading Researcher, Pavlov Institute of Physiology RAS, 199034, Saint-Petersburg, Russia; Scopus ID: 7004163777, <https://orcid.org/0000-0001-7408-302X>, vmbond@gmail.com