

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ИНВАРИАНТНОГО ВОСПРИЯТИЯ ВЕЙВЛЕТНЫХ ИЗОБРАЖЕНИЙ

© 2011 г. В. Н. Чихман, канд. техн. наук; Ю. Е. Шелепин, доктор мед. наук; С. В. Пронин

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

E-mail: niv@pavlov.infran.ru

Проведены психофизические исследования зрительного восприятия неполных контурных изображений. Изображения синтезировали с помощью вейвлетов. В качестве вейвлетов использовали DoG-функцию (Difference of Gaussians) и ориентированные вейвлеты, спектр которых ограничен как по частотному диапазону, так и по диапазону ориентаций. Варьировали размер изображений, число и угловой размер вейвлетов. Установлена зависимость между характеристиками вейвлетов и порогами распознавания неполных изображений.

Ключевые слова: неполные изображения, инвариантное зрительное восприятие.

Коды OCIS: 110.7410.

Поступила в редакцию 04.08.2011.

Введение

Зрительная система человека, наряду с возможностью восприятия индивидуальных особенностей изображений наблюдаемого объекта, обладает способностью инвариантного восприятия образов при изменении различных параметров зрительного представления объектов – освещенности, размера, ориентации. Однако до сих пор не в полной мере определены диапазоны инвариантности зрительного восприятия по широкому спектру варьируемых параметров наблюдаемых объектов. Психофизические и электрофизиологические исследования инвариантности распознавания человеком зрительных образов, помимо фундаментальных целей, имеют прикладное значение для разработки систем машинного зрения. Именно решение практических задач требует определения количественных диапазонов инвариантности [1]. Отсутствие данных о диапазонах инвариантности зрительной системы человека к различным преобразованиям параметров наблюдаемых объектов вынуждает инженеров задавать избыточные характеристики для искусственных распознающих систем.

Вопросы изучения инвариантных механизмов зрительного восприятия привлекают внимание многих исследователей. Известны работы по инвариантному описанию зрительных изображений объектов на базе так называемого

лог-полярного картирования [2–4]. Последнее характеризуется учетом геометрии проекции сетчатки в зрительную кору, при этом ретинотопика коры представлена сложным логарифмическим отображением зрительной сцены. Описание отдельных объектов в этой сцене дополняет модель согласованной фильтрации [5]. Для описания изображений и сцен применяют пирамидальное представление обработки зрительной информации [6–9]. Пирамидальное представление формируется путем последовательной свертки изображения с неким ядром (например гауссианом), причем для формирования n -го уровня пирамиды масштаб этого ядра берется пропорциональным 2^n , а частота дискретизации – обратно пропорциональной 2^n .

В данной работе стояла задача исследования в психофизиологических экспериментах инвариантности к угловым размерам зрительного восприятия неполных контурных изображений объектов.

Для организации экспериментального измерения диапазонов инвариантности зрительного восприятия необходимо было выбрать определенную количественную характеристику, которая описывает результат восприятия. В наших предыдущих работах [10–13] представлены серии экспериментов по изучению инвариантности зрительного восприятия, в которых в качестве такой характеристики использовался ре-

зультат голлин-теста [14] – теста восприятия неполных изображений. Наши предыдущие исследования инвариантности с использованием методики голлин-теста [11, 15] показали, что в диапазоне угловых размеров от 1° до 50° пороги восприятия неполных изображений не зависят от их углового размера. Однако у стандартной методики голлин-теста имеется недостаток, который можно определить как несогласованность воздействия используемых зрительных стимулов с организацией рецептивных полей зрительной системы. Основной источник этой несогласованности связан с тем, что контурные бинарные зрительные стимулы характеризуются широким спектром пространственных частот. Однако известно, что нейроны зрительной коры головного мозга реагируют на стимулы, спектры которых лежат в ограниченных диапазонах пространственных частот. Для рецептивных полей нейронов первичной зрительной коры средняя ширина этого диапазона, по данным экспериментов, равна 1,4 октавы [16–18]. Кроме того, большая часть рецептивных полей обладает ориентационной избирательностью и реагирует на стимулы, спектральные компоненты которых имеют определенный диапазон ориентаций. Спектр бинарных фрагментированных изображений стандартного голлин-теста перекрывает полосы пропускания рецептивных полей большинства нейронов, так как в нем содержатся низкочастотные и высокочастотные составляющие. Возникает вопрос – можно ли сформировать оптимальный зрительный стимул, избирательно воздействующий только на небольшую группу нейронов с близкими пространственно-частотными характеристиками? Очевидно, такой стимул должен иметь угловой размер, не превышающий углового размера рецептивных полей этой группы нейронов, а его спектр должен лежать внутри области пространственных частот, на которую эти нейроны настроены. Кроме того, экспериментально было показано, что для рецептивных полей нейронов первичной зрительной коры ширина полосы пропускания, выраженная в октавах, не зависит от пространственной частоты, на которой они дают максимальный отклик. Следовательно, спектры зрительных стимулов тоже должны обладать этим свойством, т. е. ширина их спектров в октавах не должна зависеть от частоты максимума их спектра. Этим требованиям отвечают вейвлеты – функции, имеющие хорошую локализацию как в пространственной, так

и в частотной областях. В общем случае двумерный вейвлет может быть описан как функция следующего вида:

$$f(x, y) = \frac{1}{S} g\left(\frac{x-x_0}{S}, \frac{y-y_0}{S}\right),$$

где $g(x, y)$ – функция с нулевым средним значением, отвечающая требованию хорошей пространственной и частотной локализации; S – масштаб (размер вейвлета), x_0, y_0 – координаты центра вейвлета.

Простейшим примером вейвлета является так называемая DoG-функция (Difference of Gaussians) – вейвлет, представляющий собой разность двух двумерных функций Гаусса с различной полушириной:

$$f(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) - \frac{1}{2\pi N^2\sigma^2} \exp\left(-\frac{x^2 + y^2}{2N^2\sigma^2}\right),$$

где σ – стандартное отклонение, N – масштабный коэффициент.

На рис. 1 показаны примеры DoG-функций трех размеров и их амплитудные спектры. Амплитудный спектр фрагментированного изображения, составленного из таких вейвлетов, будет лежать в том же частотном диапазоне, что и спектр одиночного вейвлета. Однако DoG-функция характеризуется тем, что ее спектр содержит спектральные составляющие со всеми возможными ориентациями. Между тем нейроны зрительной коры могут обладать ориентационной избирательностью и для экспериментов с ними может потребоваться стимул, селективно воздействующий только на клетки, чувствительные к какой-то определенной ориентации. Существуют вейвлеты, которые удовлетворяют этим требованиям, – ориентированные вейвлеты. К ним относятся, например, вейвлеты, обладающие селективностью как по пространственной частоте, так и по ориентации функции Габора, и представляющие собой синусоиду, умноженную на функцию Гаусса, –

$$f(x, y) = g(x, y) \sin(\omega(x \cos(\theta) + y \sin(\theta)) + \varphi),$$

где $g(x, y)$ – функция Гаусса, ω – пространственная частота, θ – ориентация, φ – фаза.

На рис. 2 показаны примеры ориентированных вейвлетов разных размеров и их амплитудные спектры. Итак, формируя тестовые неполные изображения из вейвлетов, можно

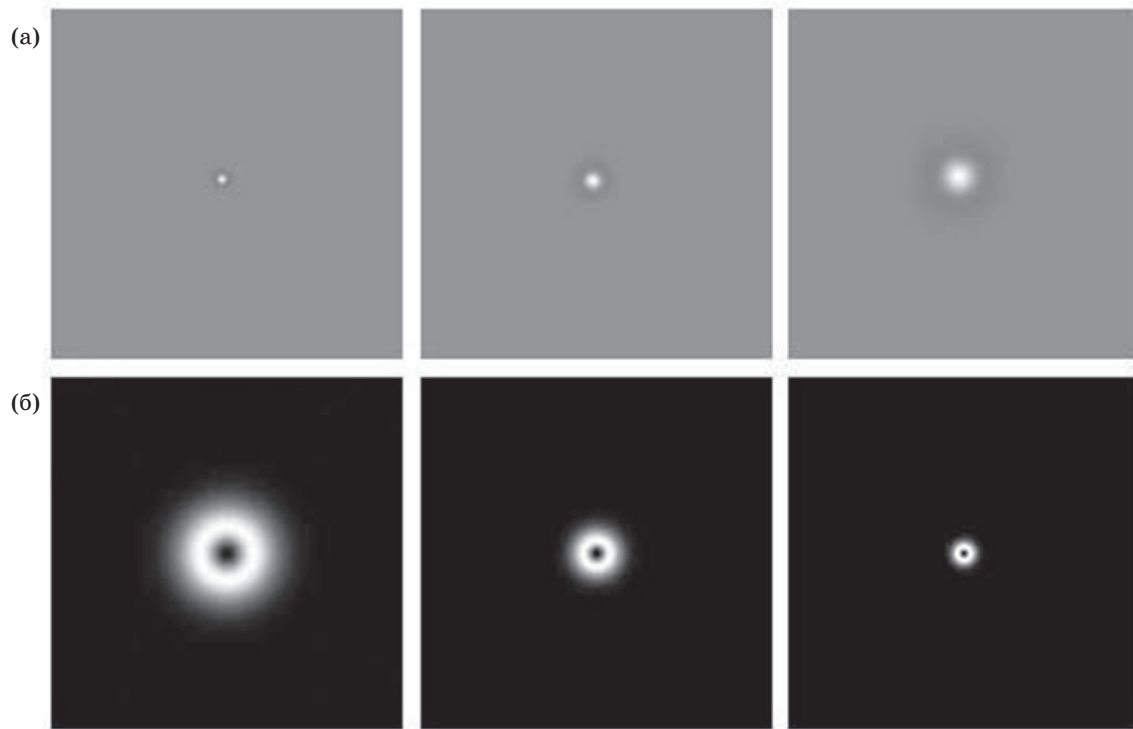


Рис. 1. Изображения DoG-функций (а) и их спектров (б).

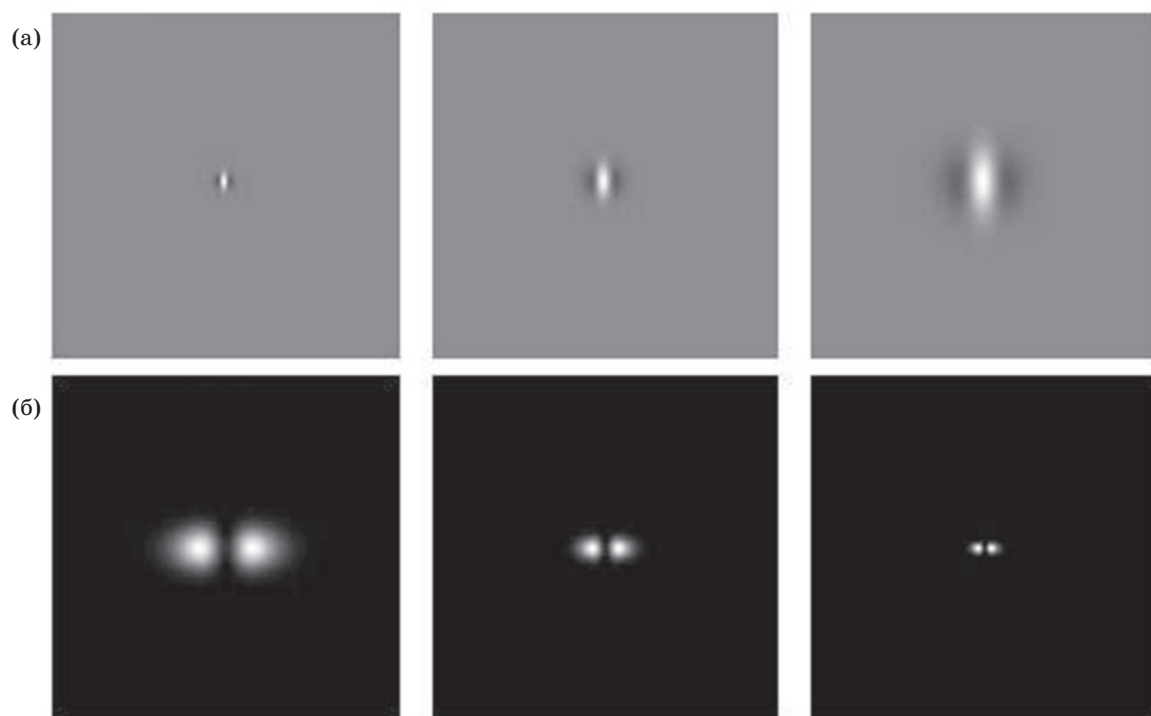


Рис. 2. Изображения ориентированных вейвлетов (а) и их спектров (б).

предполагать их воздействие только на небольшую группу зрительных нейронов с близкими пространственно-частотными полосами пропускания. К сожалению, есть один фактор, который не позволяет легко достичь та-

кой точной избирательной стимуляции. Дело в том, что спектры многих вейвлетов, в частности DoG-функций, несмотря на выраженный пик, достаточно широки и могут существенно перекрываться, что снижает избирательность

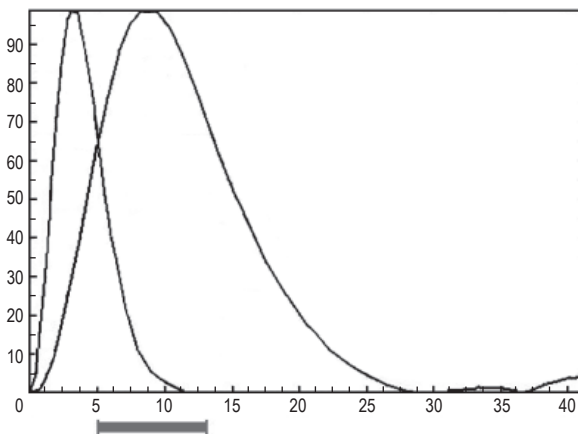


Рис. 3. Сечения двумерных спектров DoG-функций, отличающихся по угловому размеру в 2,8 раза; внизу показана ширина полосы пропускания типичной клетки первичной зрительной коры, равная 1,4 октавы.

воздействия сформированных на их основе визуальных стимулов. В качестве иллюстрации на рис. 3 показаны сечения двумерных спектров двух DoG-функций, отличающихся по угловому размеру в 2,8 раза.

Видно, что их спектры существенно перекрываются, хотя их максимумы хорошо разделены. Для сравнения внизу показана ширина полосы пропускания типичной клетки первичной зрительной коры, равная 1,4 октавы. Очевидно, в эту полосу попадают оба спектра. В результате на вейвлет, который является оптимальным стимулом для определенной группы нейронов, будут реагировать также и другие нейроны, если пространственно-частотные диапазоны их рецептивных полей даже незначительно перекрываются со спектром данного вейвлета. Чтобы устранить этот нежелательный эффект, можно использовать простой прием понижения контраста вейвлетов до порогового уровня. Основанием для этого служит предположение, что при пороговом контрасте вейвлета на него будут реагировать только те нейроны, для которых он является оптимальным стимулом. У нейронов, для которых этот вейвлет не является оптимальным стимулом, реакция на стимул окажется ниже уровня внутреннего шума. Таким образом, можно дискриминировать избирательное воздействие стимулов по пространственной частоте, т. е. стимулировать только те нейроны, для которых данный стимул является близким к оптимальному.

Методика

Психофизический эксперимент с использованием стандартной методики голлин-теста проводится следующим образом. Имеется набор из нескольких десятков контурных изображений общеизвестных объектов, хорошо знакомых испытуемому. Компьютерная программа разбивает контуры этих изображений на фрагменты фиксированного размера, которые выводятся в случайном порядке на экран монитора, постепенно формируя полный контур объекта. Вывод фрагментов останавливается в момент, когда испытуемый распознал изображение. При этом программа фиксирует в момент распознавания степень пороговой фрагментации изображения, которая равна отношению суммарной площади предъявленных фрагментов контура к полной площади контура. Чем выше этот процент, тем хуже распознается изображение.

В проведенных нами измерениях испытуемым предъявлялись в режиме Голлин-теста неполные изображения, состоящие из цепочек вейвлетов. Для формирования тестовых изображений использовались как DoG-функции, так и ориентированные вейвлеты. В качестве тестовых изображений-стимулов были использованы контуры букв и цифр (всего 24 изображения). При синтезе неполного изображения контур буквы или цифры аппроксимировался цепочкой вейвлетов с заданным постоянным шагом между ними. При проведении измерений программа в случайном порядке выводила вейвлеты на экран, постепенно формируя контур объекта. Пороговый уровень фрагментации фиксировался в тот момент, когда испытуемый распознавал изображение. Порог восприятия определялся как процент вейвлетов, выведенных на экран в момент распознавания объекта испытуемым, от общего числа вейвлетов в синтезированном неполном изображении на выбранном уровне пирамидального представления.

В измерениях участвовало 5 испытуемых.

Результаты и обсуждение

Были проведены измерения по исследованию инвариантности зрительного восприятия к угловому размеру изображений. Угловым размером изображений менялся от 1° до 4° . С каждым испытуемым было проведено несколько серий измерений с единичным и с пороговым

контрастом. Под пороговым контрастом понимался минимальный контраст, при котором испытуемый мог обнаружить на экране изображение вейвлета. Этот пороговый контраст определялся отдельно для каждого испытуемого и для каждого размера вейвлета. Пример использованных фрагментированных изображений, составленных из ориентированных вейвлетов, а также часть полученных результатов показаны на рис. 4. Видно, что при единичном контрасте инвариантность к угловому размеру в данных пределах сохраняется, что согласуется с полученными нами ранее данными [13, 15]. Таким образом, при надпороговом контрасте ответ инвариантен относительно размера изображения. Однако из рисунка видно, что в случае порогового контраста вейвлетов инвариантность восприятия отсутствует: чем больше угловой размер изображений, тем хуже они распознаются. Возникает вопрос – в чем причина этого нарушения инвариантности? Предлагается следующее объяснение этих результатов. По мере увеличения

углового размера изображения оно выходит за пределы фовеальной области. Но чем больше угловое расстояние от центра фовеальной области, тем (в среднем) больше размер рецептивных полей нейронов, причем максимум их чувствительности смещается в область низких пространственных частот [18–20]. В результате небольшие по размеру вейвлеты, из которых состоят данные изображения, не являются оптимальными стимулами для этих нейронов. И если контраст вейвлетов близок к пороговому, нейроны на такие стимулы не реагируют из-за высокого уровня внутреннего шума. Если вейвлеты небольшие, испытуемый рассматривает изображение как бы через узкое отверстие, через которое виден лишь его небольшой фрагмент. Естественно, в этих условиях распознавание ухудшается. Следует заметить, что в этом эксперименте при изменении углового размера изображений угловой размер самих вейвлетов оставался одинаковым в отличие от ситуации, когда тестовые изображения рассматриваются с различных дистанций (в последнем случае с изменением углового размера фигур пропорционально менялся бы и размер вейвлетов).

В другой серии измерений было исследовано влияние на процесс распознавания изменения размеров самих вейвлетов при прочих равных условиях. Угловой размер всех изображений был равен двум угловым градусам, следовательно, изображения целиком попадали в фовеальную область. Пример использованных в этих экспериментах изображений и полученные результаты иллюстрирует рис. 5. На графике по оси ординат отложена доля предъявленных вейвлетов, при которой произошло распознавание. Первая диаграмма показывает результаты измерений при размере вейвлетов 5,4'. Из диаграммы видно, что при единичном контрасте порог был равен 73% (столбик 1), при пороговом контрасте – 96% (столбик 2). Это означает, что большинство неполных изображений при пороговом контрасте вообще не распознавались. Они распознавались только тогда, когда почти весь контур был выведен на экран. Вторая диаграмма – те же измерения, но при размере вейвлетов в 21,6'. Видно, что для единичного контраста результаты практически не изменились (столбик 1), в то время как для порогового контраста неполные изображения стали уверенно распознаваться и порог понизился до 73,3% (столбик 2). То есть увеличение угловых размеров вейвлетов, при

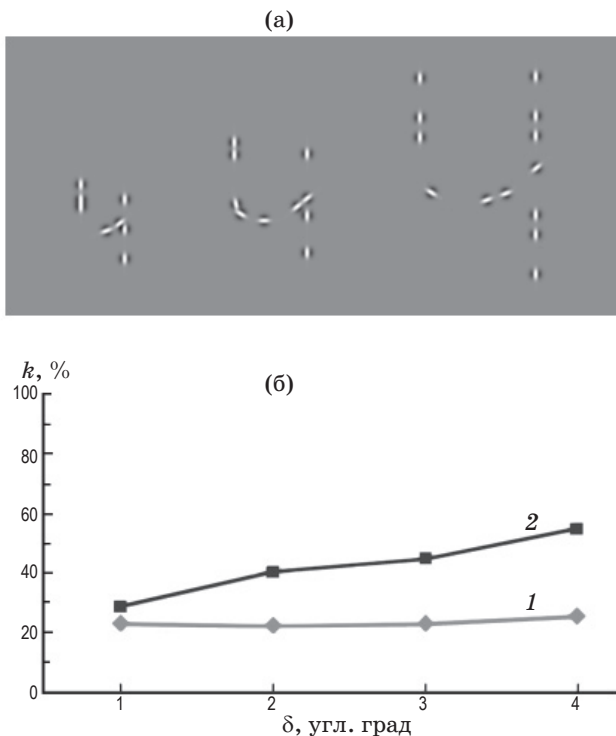


Рис. 4. а – примеры фрагментированных изображений разного размера из цепочек одинаковых ориентированных вейвлетов; б – зависимости порога восприятия (доля k от общего числа вейвлетов, при котором произошло распознавание) от углового размера изображения δ при единичном (1) и пороговом (2) контрасте.

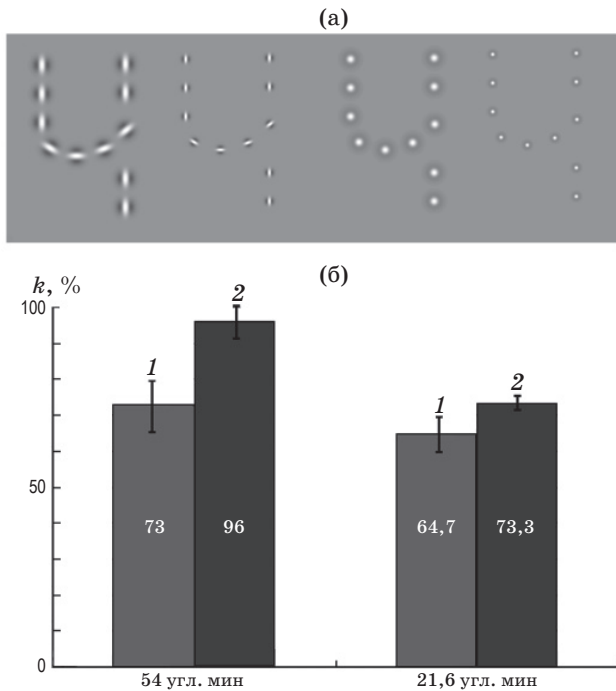


Рис. 5. а – примеры неполных изображений одного размера из цепочек вейвлетов разного размера; б – диаграммы зависимости доли k предъявленных вейвлетов, при которой произошло распознавание. На оси абсцисс указаны размеры вейвлетов. 1 – при единичном контрасте, 2 – при пороговом контрасте.

прочих равных условиях, улучшило распознавание тестовых изображений.

Можно предложить следующее объяснение результатов этих измерений. В процессе распознавания объекта происходит объединение информации от клеток с рецептивными полями, расположенными в различных местах поля зрения; при этом можно предположить, что на ранних этапах обработки процесс анализа локализован, т. е. объединяется информация только с тех рецептивных полей, которые расположены в поле зрения рядом друг с другом, не дальше, чем некое максимальное расстояние D в поле зрения. Если угловая дистанция между вейвлетами во фрагментированном изображении больше D , процесс объединения отдельных вейвлетов в единый контур затруднен, что приводит к ухудшению распознавания. Предположим, что эта дистанция D в угловых еди-

ницах тем больше, чем больше угловой размер рецептивного поля. Увеличение углового размера вейвлетов в данных экспериментах приводило к тому, что они становились оптимальными стимулами для нейронов с более крупными угловыми размерами рецептивных полей. Объединение информации от этих полей, в соответствии с пирамидальным представлением информации, происходит на больших угловых расстояниях, что и улучшало распознавание при меньшем числе выведенных вейвлетов. Следует заметить, что если рассматривать величину внутрикорковых связей в миллиметрах, то в этом случае внутрикорковые связи, вероятно, остаются постоянными по длине как в корковом представительстве центра поля зрения, так и на периферии [21, 22]. Таким образом, полученные данные косвенно подтверждают представление о лог-полярном кодировании в зрительной коре.

Выводы

1. При надпороговом контрасте восприятие инвариантно к преобразованию размера изображения, что подтверждает наши предыдущие результаты и может быть объяснено высоким отношением сигнал/шум при единичном контрасте стимула.

2. Показано, что в случае порогового контраста вейвлетов инвариантность восприятия отсутствует, причем чем больше угловой размер изображений, тем хуже они распознаются. Это можно объяснить как результат выхода изображений за пределы фовеальной области, где вейвлеты, из которых они состоят, уже не являются оптимальными стимулами. Реакция на стимул порогового контраста оказывается ниже уровня внутреннего шума.

3. Увеличение размеров вейвлетов при заданном размере изображения приводит к улучшению восприятия при пороговом контрасте, предположительно, вследствие того, что взаимодействие рецептивных полей происходит на угловых расстояниях, пропорциональных угловым размерам этих полей.

Работа поддержана грантом РФФИ 09-07-00336.

* * * * *

ЛИТЕРАТУРА

1. Ghosh A., Petkov N. Robustness of shape descriptors to incomplete contour representations // IEEE Transactions on pattern analysis and machine intelligence. 2005. V. 27. № 11. P. 1793–1804.

2. *Mehanian C., Rak S.* Bidirectional log-polar mapping for invariant object recognition // Proc. SPIE. 1991. № 1471. P. 200.
3. *Weiman C.F.R.* Log-polar vision for mobile robot navigation // Electronic Imaging. 1990. P. 382–385.
4. *Zokai S., Wolberg G.* Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations // IEEE Transactions on Image Processing. 2005. V. 14. № 10. P. 1422–1434.
5. *Красильников Н.Н., Шелепин Ю.Е.* Функциональная модель зрения // Оптический журнал. 1997. Т. 64. № 2. С. 72–82.
6. *Александров В.В., Горский Н.Д.* Представление и обработка изображений. Рекурсивный подход. Л.: Наука, 1985. 190 с.
7. *Шелепин Ю.Е., Бондарко В.М., Данилова М.В.* Конструкция фовеолы и модель пирамидальной организации зрительной системы // Сенсорные системы. 1995. Т. 9. № 1. С. 87–97.
8. *Burt P., Adelson E.* The Laplacian Pyramid as a Compact Image Code // IEEE Transactions on communications. 1983. V. Com-31. № 4. P. 532–540.
9. *Cantoni V., Petrosino A.* Neural Recognition in a Pyramidal Structure // IEEE Transactions on neural networks. 2002. V. 13. № 2. P. 472–480.
10. *Shelepin Y., Vahromeeva O., Harauzov A., Pronin S., Foreman N., Chihman V.* Recognition of incomplete contour and half-tone figures // Perception. 2004. V. 33. Supplement. P. 85.
11. *Chikhman V., Shelepin Y., Foreman N., Merkuljev A., Pronin S.* Incomplete figure perception and invisible masking // Perception. 2006. V. 35. № 11. P. 1441–1457.
12. *Шелепин Ю.Е., Чихман В.Н., Фореман Н.* Анализ исследований восприятия фрагментированных изображений: целостное восприятие и восприятие по локальным признакам // Российский физиологический журнал. 2008. Т. 94. № 7. С. 758–776.
13. *Шелепин Ю.Е., Чихман В.Н., Вахрамеева О.А., Пронин С.В., Фореман Н., Пэмор П.* Инвариантность зрительного восприятия // Экспериментальная психология. 2008. № 1. С. 7–33.
14. *Gollin E.* Developmental studies of visual recognition of incomplete objects // Perceptual and Motor Skills. 1960. № 11. P. 289–298.
15. *Вахрамеева О.А., Шелепин Ю.Е., Мезенцев А.Ю., Пронин С.В.* Изучение восприятия неполных контурных изображений различного размера // Российский физиологический журнал. 2008. Т. 94. № 10. С. 1158–1170.
16. *DeValois R.L., Albrecht D.G., Thorell L.G.* Spatial frequency selectivity of cells in macaque visual cortex // Vision Res. 1982. V. 22. P. 545–559.
17. *Шелепин Ю.Е.* Фильтрационные свойства рецептивных полей нейронов зрительной коры // Докл. АН СССР. 1981. Т. 261. № 6. С. 1506–1509.
18. *Шелепин Ю.Е.* Пространственно-частотные характеристики рецептивных полей нейронов латеральной супрасильвиевой области // Нейрофизиология. 1982. Т. 14. № 6. С. 608–614.
19. *Ginsburg A.* Spatial filtering and visual form perception // In Handbook of Perception and Human Performance /Ed. K. Boff. New York: John Wiley and Sons. 1986. V. 34. P. 1–41.
20. *Шелепин Ю.Е., Колесникова Л.Н., Левкович Ю.И.* Визоконтрастометрия. Измерение пространственных передаточных функций зрительной системы. Л.: Наука, 1985. 105 с.
21. *Алексеев С.В., Топорова С.Н., Макаров Ф.Н.* Микротопография корковых полей 17 и 18 // Сенсорные системы. 1999. Т. 13. № 4. С. 278–283.
22. *Gilbert C.D.* Microcircuitry of the visual cortex // Ann. Rev. Neurosci. 1983. V. 6. P. 217–247.