

ИССЛЕДОВАНИЕ РОЛИ ПРОСТРАНСТВЕННО-ЧАСТОТНОЙ ФИЛЬТРАЦИИ ИЗОБРАЖЕНИЙ ПРИ ОЦЕНКЕ ВОЗРАСТА И ВОСПРИЯТИЯ ЭМОЦИОНАЛЬНОГО ВЫРАЖЕНИЯ ЛИЦ

© 2015 г. Е. В. Логунова^{*}; Ю. Е. Шелепин^{**}, доктор медицинских наук

^{*}Санкт-Петербургский государственный университет, Санкт-Петербург

^{**}Университет ИТМО, Санкт-Петербург

E-mail: yshelepin@yandex.ru

В работе исследован процесс восприятия сцен с изображением лица человека, подвергнутых обработке фильтрами пространственных частот, имитирующими характеристики рецептивных полей нейронов первичной зрительной коры. Была использована методика, позволяющая дать количественную оценку восприятия изображения при выполнении задач идентификации выявления периода эмоционального напряжения и возрастных особенностей человека. Было показано, что в процессе восприятия изображений лиц, помимо горизонтальных составляющих пространственно-частотного спектра, существенную роль играют диагональные составляющие. Несмотря на то, что чувствительность зрительной системы к диагональным составляющим меньше, чем к горизонтальным, содержащаяся в них информация позволяет не только выделять возрастные особенности, но и давать дополнительную информацию, необходимую для идентификации незнакомого человека при повторной встрече.

Ключевые слова: зрение, распознавание, лицо, пространственно-частотная фильтрация, идентификация, возраст, эмоциональное выражение, ориентационная избирательность.

Коды OCIS: 100.4999, 100.6950, 330.5000.

Поступила в редакцию 05.05.2015.

Введение

Конструкция “зрительного мозга” такова, что при построении образа сложной сцены зрительная система человека анализирует физические параметры воспринимаемого сигнала и выбирает из них значимые для наблюдателя свойства.

Известно [1–6], что для распознавания большого класса объектов достаточно только вертикальных и горизонтальных составляющих пространственно-частотного спектра изображений. На использовании только вертикальных и горизонтальных составляющих основан широко применяемый в области технического зрения метод детектирования объектов Виолы–Джонса [7]. С. Дакин и Р. Ватт [1] при цифровой обработке статичных изображений лиц показали, что горизонтальные составляющие позволяют эффективно выделять лицо и используются при идентификации знакомого человека. Однако

при анализе потока сцен, где основной задачей является выделение фигуры человека, зрительная система в первую очередь выстраивает шаблон для движущегося объекта преимущественно за счет вертикальных составляющих. Это и позволяет при восприятии соотнести наблюдаемый объект с определенной семантической категорией, после чего с помощью горизонтальных составляющих строится “штрих-код лица”, на основе которого и происходит идентификация человека [1, 2]. Однако для выявления паттернов невербальной коммуникации, проявляющихся в минимальных изменениях мимики лица, уже недостаточно информации, содержащейся в горизонтальных и вертикальных пространственно-частотных составляющих изображений. Диагональные составляющие позволяют выделить незначительную, практически скрытую информацию. Такое выделение происходит неосознанно в рецептивных полях затылочной коры в результате процессов, напо-

минающих пространственно-частотную фильтрацию с использованием вейвлетов, подобных функциональным элементам Габора. Эту фильтрацию можно рассматривать как особую маску, “грим”, подчеркивающий мимические особенности при восприятии лиц, которые являются опорными сигналами для механизмов принятия решения на более высоких уровнях зрительной системы [8]. Однако остается актуальным вопрос о роли пространственно-частотных фильтров при решении задачи “незнакомое лицо”, кто он, “свой–чужой”, “он–она”, “возраст”, “эмоциональное состояние”. Остается неизученным процесс идентификации одного и того же человека с учетом изменения ситуативных особенностей, например возрастных изменений или эмоционального напряжения.

Одним из методов для решения поднятых проблем является моделирование. В данном случае метод является не только экономичным по временным показателям и затраченным ресурсам, но и прогрессивным способом решения поставленных задач. Известно, что подобное моделирование связано с различными способами фильтрации [5–8]. Использование методов пространственно-частотной фильтрации позволяет имитировать работу ориентационно-избирательных рецептивных полей затылочной коры, настроенных на выделение горизонтальных и вертикальных составляющих в пространственно-частотном спектре наблюдаемого изображения.

Профили весовых функций рецептивных полей корковых нейронов хорошо описываются синусоидой, промоделированной гауссианом, т.е. элементом Габора [3–6], поэтому фильтры, основанные на данных элементах, эффективно выделяют объекты в сложной сцене. Цель данной работы – исследование возможностей пространственно-частотной фильтрации наблюдаемых изображений для предварительной оценки незнакомого человека с учетом его возраста и эмоционального состояния.

Методика исследования

В исследованиях в ходе проведения каждого эксперимента приняло участие по 34 человека, из которых 47% женщин и 53% мужчин в возрасте 25–45 лет.

Разработанная нами методика используется для моделирования и оценки особенностей информации, получаемой от определенных про-

странственно-частотных характеристик при восприятии сцены “Лицо”.

При подборе тестовых стимулов для изучения роли различных фильтров при восприятии сцены “Незнакомое лицо” было определено три следующих условия:

- 1) воспринимаемое лицо неизвестно воспринимающему его человеку,
- 2) воспринимаемое лицо человека отображено в различные периоды жизни,
- 3) один из периодов отличается эмоциональной напряженностью.

В качестве стимулов методики были взяты фотографии Клер Фелиси из проекта “Here are the Young Men (Marked)” [9]. Это черно-белые триптих-портреты-снимки, объединенные по три изображения, относящиеся к трем периодам жизни, один из которых война. В методике фотографии разделены на три группы, каждая группа характеризует один из периодов жизни, которые получили условное название – “До”, “Во время”, “После”. Война является незабываемым и сильно влияющим фактором в жизни человека, она может поменять не только характер и поведение, но и изменить взгляд, мимику. Кроме того, такие жизненные обстоятельства, как военные действия или катастрофы, могут более четко оставлять “возрастной отпечаток” и давать возможность отследить влияние эмоциональных нагрузок и стресса при восприятии сцены “Лицо”. Выбранные фотографии (всего было отобрано 11 триптих-портретов) относятся к категории “Незнакомое лицо”.

Для сравнения особенностей восприятия сцен “Незнакомое лицо” и “Знакомое лицо” в качестве стимулов предъявлялись сцены с изображением лиц популярных актеров-мужчин (использовалось три варианта изображения лица одного и того же актера в разные периоды жизни).

Для моделирования работы рецептивных полей изображения были подвергнуты цифровой фильтрации [2–6] путем свертки с элементом Габора, которая имитирует работу рецептивных полей нейронов зрительной коры. Первый вариант обработки изображения лиц – горизонтальная фильтрация, второй вариант – диагональная. Предыдущие исследования [2] показали, что вертикальная фильтрация эффективна только для выделения фигуры человека на сложном фоне, а для восприятия лица является неактуальной. Результат такой цифровой фильтрации – это некий аналог схемы изображения,

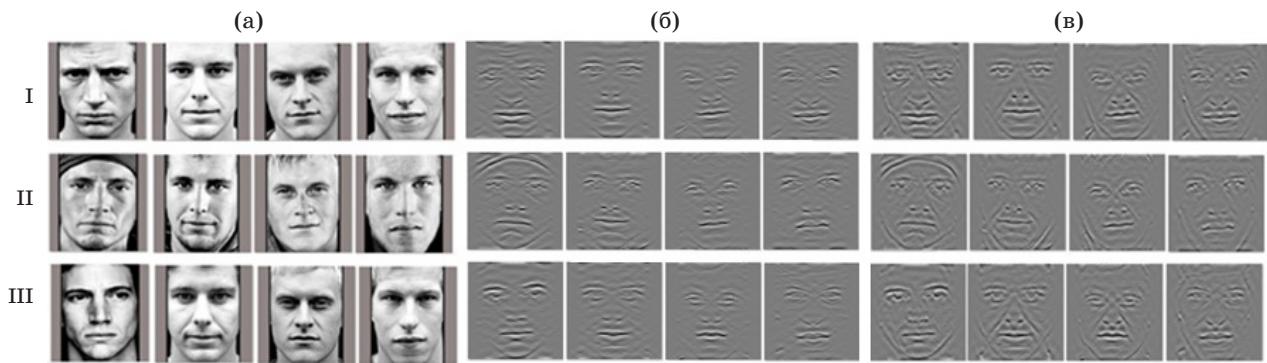


Рис. 1. Сцены “Незнакомое лицо”, используемые в экспериментах. Стока I – изображения из группы периода “До”; строка II – изображения из группы периода “Во время”; строка III – изображения из группы периода “После”; столбцы под буквой а) – исходные фотографии; столбцы под буквой б) и в) – примеры изображений после горизонтальной и вертикальной вейвлетной фильтрации соответственно.

которая сравнивается с динамическим шаблоном, хранящимся в памяти человека [10, 11]. Примеры сцен из категории “Незнакомое лицо”, используемые в методике, приведены на рис. 1.

Для оценки характера получаемой информации используется последовательное предъявление стимулов и их последующее опознание в группе похожих сцен либо их категоризация в зависимости от изучаемого фактора. В ходе исследования было проведено три эксперимента, в каждом эксперименте выбранные сцены предъявлялись в три этапа. На первом этапе использовались исходные фотографии (рис. 1а), на втором и третьем – изображения после вертикальной и горизонтальной вейвлетной фильтрации соответственно (рис. 1б и 1в). Перерыв между этапами 1,5–2 недели. На каждом из этапов использовалось по 33 изображения.

Эксперимент № 1

Для выявления характера информации, содержащейся в горизонтальных и диагональных пространственно-частотных составляющих, при восприятии сцены “Лицо” в процессах “идентификация человеком” и “идентификация человека с учетом ситуативных особенностей” был проведен первый эксперимент.

На короткое время (2,5 с) предъявляли одну из сцен категории “Незнакомое лицо”, которую предлагалось запомнить. После этого из 33 исходных фотографий следовало выбрать лицо того человека, которое предъявлялось до этого. В ходе тестирования использовалось три изображения лица одного и того же человека, изображение варьировалось в зависимости от пери-

ода жизни. Схема прохождения эксперимента показана на рис. 2.

После предъявления сцен из категории “Незнакомое лицо” таким же образом респондентам предлагалось оценить изображения сцен из категории “Знакомое лицо”. В данном исследовании процессы “категоризации объекта как лицо” и “идентификация лица с конкретным человеком” объединены и условно названы задачей “Узнавание”.

При анализе результатов первого эксперимента вычислялось среднее значение доли верных ответов по группе в целом. Значение доли рассчитывалось как соотношение количества

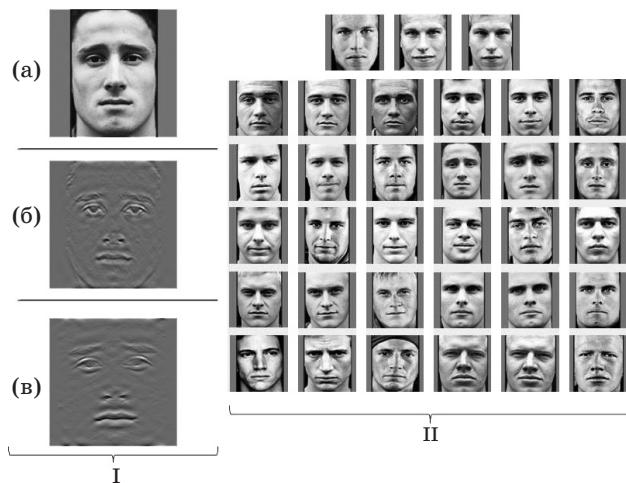


Рис. 2. Схема проведения первого эксперимента. Стока I – вариант изображения стимула для дальнейшей идентификации; строка II – группа стимулов для выбора предъявленного ранее; (а), (б), (в) – стимулы первого, второго, третьего этапов соответственно.

правильных ответов к общему количеству предъявляемых сцен. В зависимости от изучаемого процесса учет количества правильных ответов изменялся следующим образом:

1) для задачи “Узнавание” правильными ответами считались все ответы, в которых респондент правильно указывал человека вне зависимости от периода его жизни,

2) для задачи “Идентификация человека в конкретной ситуации” (условное обозначение “Ситуация”) правильными ответами являются все ответы, в которых респондент правильно указывал человека с учетом конкретной жизненной ситуации.

Интервальная оценка доли выбранных изображений при доверительном уровне 0,9 рассчитана с использованием квантиля распределения Стьюдента, так как дисперсия оцениваемого параметра предполагается неизвестной. Такой параметр позволяет показать динамику оценки не только в данной группе, но и прогнозировать результаты оценки анализируемых процессов при повторных экспериментах. Статистическая достоверность различия была подтверждена с использованием двустороннего критерия сравнения двух выборочных долей.

Эксперименты № 2 и № 3

Для выявления роли пространственно-частотных составляющих в процессах “выявление периода эмоционального напряжения (стресс)” и “определение возрастных особенностей” были проведены второй и третий эксперименты.

В процессе тестирования респонденту одновременно предъявляли три сцены изображения лица одного и того же человека, относящиеся к разным периодам жизни. При проведении второго эксперимента из них предлагалось выбрать изображение, на котором наиболее ярко выражен стрессовый период. При проведении третьего эксперимента предлагалось выбрать изображение, относящееся к самому поздне-

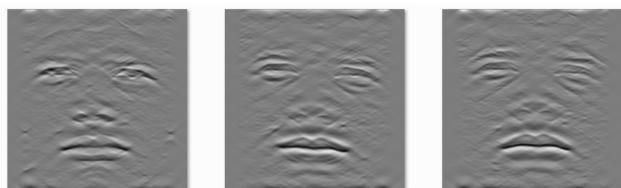


Рис 3. Пример предъявления стимулов при проведении второго и третьего эксперимента.

му периоду жизни. Пример одной из итераций предъявления стимулов при проведении второго и третьего экспериментов приведен на рис. 3.

Для их проведения использовались изображения, относящиеся к категории “Незнакомое лицо”. Этапы, периодичность и особенности предъявления стимулов во втором и третьем экспериментах соответствуют первому.

При анализе результатов этих экспериментов вычислялось среднее значение доли выбранных изображений для каждого из трех периодов: “До”, “Во время”, “После”. Такое разделение сцен позволяет проследить влияние факторов “Возраст” и “Стресс” при восприятии сцен “Лицо”. Значение доли рассчитывалось как отношение количества выбранных изображений, соответствующих определенному периоду жизни, к общему количеству предъявляемых сцен.

Результаты и их обсуждение

Результаты эксперимента № 1

Дакин С. и Ватт Р. в своих исследованиях показали, что горизонтальных составляющих достаточно для эффективного выполнения следующих задач: “категоризация – определение, что объект – это лицо человека” и “идентификация – соотнесение воспринимаемого лица с конкретным человеком”, назвав данный эффект “штрих-кодом”, необходимым для идентификации конкретного, известного ранее человека [1]. При этом, несмотря на различные манипуляции со стимулами (изменение размера, освещения, ракурса), “штрих-код” остается неизменным для конкретного лица и позволяет узнать воспринимаемый образ. Однако стоит подчеркнуть, что исследовались лица популярных людей, хорошо известных воспринимающей аудиторией, т.е. объектом восприятия была сцена “Знакомое лицо”. Изменение семантического признака “Знакомое” на “Незнакомое” в процессах восприятия сцены может иметь свои нюансы. Среднее значение доли правильных ответов, описывающее динамику обоих процессов, показано на рис. 4.

Доля правильно оцененных изображений при выполнении задачи “узнавание” для сцен категории “Знакомое лицо” вне зависимости от особенностей психофизических характеристик воспринимаемых сцен высока (исходные

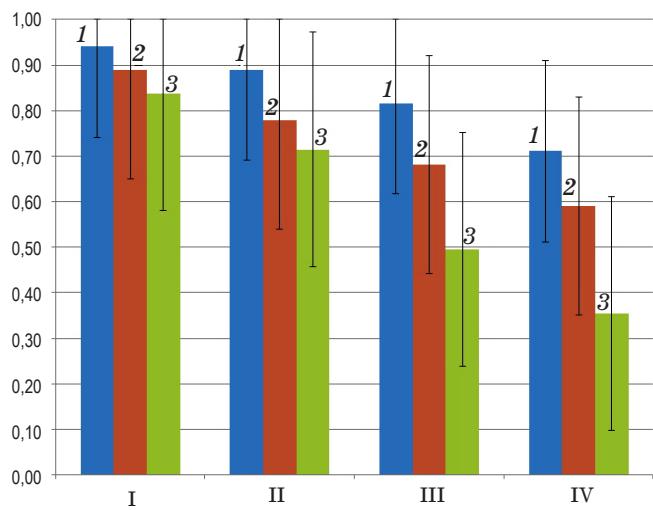


Рис. 4. Диаграмма распределения доли правильных ответов в зависимости от задач и методов фильтрации. По оси абсцисс – доля правильных ответов, по оси ординат – сцены категории “знакомое лицо”, процесс “узнавание” (I), сцены категории “знакомое лицо”, процесс “ситуация” (II), сцены категории “незнакомое лицо”, процесс “узнавание” (III), сцены категории “незнакомое лицо”, процесс “ситуация” (IV). Исходные фото (1), диагональная фильтрация (2), горизонтальная фильтрация (3).

фото – 0,94; изображения после диагональной и горизонтальной фильтрации – 0,89 и 0,84 соответственно), достоверных отличий не обнаружено. Данные результаты еще раз показывают, что горизонтальных составляющих достаточно для идентификации раннее известного человека. Динамический шаблон, сформировавшийся в памяти человека за счет неоднократного повторения, моментально активируется и дает возможность узнать человека только по “штрих-коду” [2, 3]. Кроме того, следует отметить, что достоверных отличий не обнаружено и при анализе результатов, иллюстрирующих процесс “идентификация человека с учетом ситуативных особенностей”, т.е. горизонтальных составляющих достаточно для идентификации изменившегося знакомого ранее человека.

Характер процессов идентификации изменяется при выполнении тех же задач, но для категории сцен “Незнакомое лицо” (доля правильно оцененных изображений: исходные фото – 0,82; изображения после диагональной и горизонтальной фильтрации – 0,69 и 0,48 соответственно). Достоверных отличий нет между оценками результатов, иллюстрирующих процесс “узнавание” ($p = 0,26$) и процесс “иденти-

фикация с учетом ситуативных особенностей” ($p = 0,13$), при предъявлении исходных фото с изображениями, прошедшиими диагональную фильтрацию. Однако значимые отличия наблюдаются при оценке изображений, прошедших горизонтальную фильтрацию. Горизонтальных линий становится недостаточно не только для учета ситуативных особенностей ($p = 0,0073$), но и для самого процесса идентификации ($p = 0,0049$). Дополнительная информация, имеющаяся в диагональных составляющих, помогает определить особенности формы лица, его размер, что дает определенные реперные точки, которые с большей вероятностью исключают возможность ошибок. Сложившийся в памяти человека штрих-код эффективен только для восприятия сцен “Знакомое лицо”, а анализ диагональных составляющих необходим для формирования динамического шаблона при восприятии сцен “Незнакомое лицо”. Динамика процесса “узнавание” отличается от процесса “идентификация с учетом ситуативных особенностей”. Количество правильных ответов при выполнении второй задачи снижается вне зависимости от психофизических характеристик предъявляемых стимулов и категории распознаваемых сцен. Анализ результатов показывает, что при выполнении обеих задач наблюдается одинаковая тенденция. Вне зависимости от семантического признака сцены “Знакомое” или “Незнакомое” пусковым механизмом для пространственно-частотных фильтров является процесс категоризация объекта сцены как “Лицо”.

Результаты экспериментов № 2 и № 3

Немаловажной задачей при восприятии сцен “Незнакомое лицо” является возможность диагностики стрессового поведения. Стресс является фактором эмоционального напряжения человека, а его влияние и последствия чаще всего отражаются на лице. Находясь в стрессовом состоянии, человек способен на различные поступки, в том числе и на проявление агрессии по отношению к другому человеку. Неосознанный анализ мимических особенностей человека нередко дает возможность избежать определенной опасности. Для успешной социализации человеку необходимо выполнение такой задачи, как “определение возрастных особенностей”, которые помогают выстраивать эффективные коммуникативные схемы.

Особенности распределения оценок респондентов для изучаемых факторов “Стресс” и “Возраст” с расчетом доверительного интервала по каждому из трех периодов (“До”, “Во время”, “После”) представлены на рис. 5.

При оценке сцен на наличие фактора “стресс” (I, II, III на рис. 5) было выявлено, что чаще всего наличие фактора “стресс” при оценке исходных фотографий отмечалось у сцен, относящихся к периоду военных действий (категория “Во время”), доля отмеченных изображений – 0,45. Доля изображений, относящихся к сценам из категории “После”, – 0,36, из категории “До” – 0,18. При оценке фотографий, прошедших как горизонтальную, так и диагональную фильтрацию, наличие фактора “стресс” чаще всего отмечалось у изображений, относящихся к периоду послевоенных действий (категория сцен “После”) – 0,45. Доля сцен, относящихся к категории “Во время” при горизонтальной фильтрации – 0,36, при диагональной обработке – 0,34.

Изображения сцен, прошедших как горизонтальную, так и диагональную фильтрацию, при выявлении фактора “стресс” отмечают два периода – “Во время” и “После”. Они наиболее четко указывают на постстрессовые изменения человека, что малозаметно на обычных фотографиях. Исследования показывают, что мимические изменения отражаются при горизонтальной фильтрации и служат показателем

сильного эмоционального напряжения. Это означает, что чувствительность зрительной системы к горизонтальным составляющим может являться основой невербальных сигналов опасности, идущей от незнакомого человека.

В ходе проведения третьего эксперимента (IV, V, VI на рис. 5) было выявлено, что при оценке исходных фотографий чаще всего наиболее поздний возрастной период испытуемых соотносили со сценами из категории “Во время” (0,60). Фактор “Возраста” был соотнесен с одним из самых напряженных эмоциональных периодов жизни (период военных действий), и эмоциональная нагрузка была воспринята как возрастной отпечаток. При оценке изображений, прошедших горизонтальную фильтрацию, четкого разделения не получилось. Самый поздний возрастной период соотносили как с изображениями сцен из категории “Во время”, так и из категории сцен “После” (0,36 для каждого периода). Это говорит о том, что горизонтальные составляющие не играют значительной роли при выполнении задачи “определение возраста”. Наиболее эффективными для данного процесса оказались диагональные фильтры (0,67).

Определение возраста человека является одной из сложнейших задач и в обычной жизни решается очень часто с большим количеством ошибок. Скорее всего, это объясняется тем, что задача “определение возраста человека” не является процессом, обеспечивающим безопасность

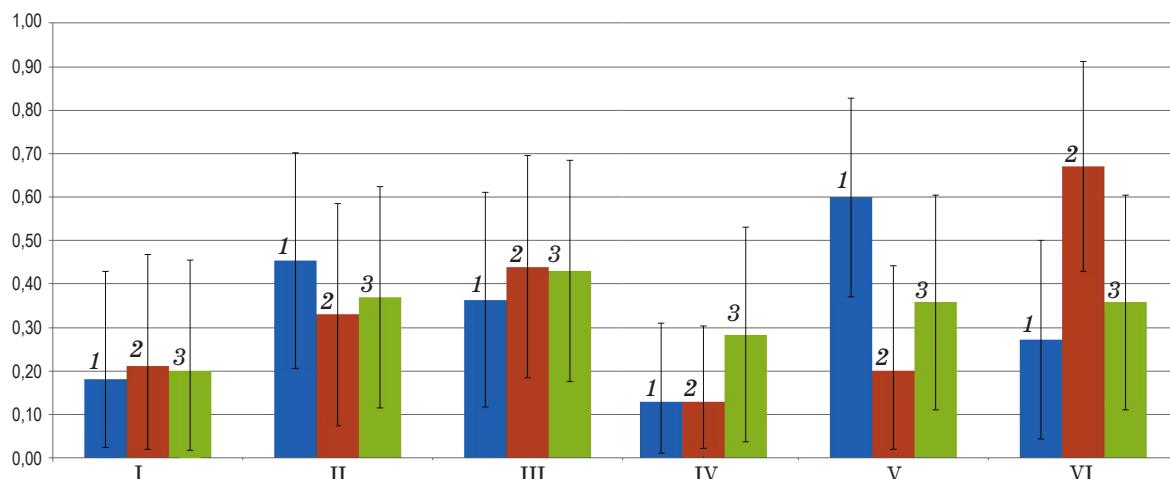


Рис. 5. Диаграмма распределения доли выбранных изображений, соотнесенных с фактором “стресс” и “возраст” в зависимости от периодов жизни и методов фильтрации. По оси абсцисс – доля отмеченных изображений, относящихся к изучаемому явлению. По оси ординат – оценки фактора “стресс” сцены категории “До” (I), сцены категории “Во время” (II), сцены категории “После” (III), сцены категории “До” (IV), сцены категории “Во время” (V), сцены категории “После” (VI). Исходные фото (1), диагональная фильтрация (2), горизонтальная фильтрация (3). Вертикальная линия –доверительный интервал доли вероятности.

и жизнедеятельность человека, а становится лишь элементом, необходимым для установления иерархических отношений в социуме, которые не всегда необходимы. В процессе социализации человека диагональная чувствительность является не только естественным свойством разделимости, которая приводит к уменьшению вычислительной сложности воспринимаемой сцены, она подчеркивает возрастные особенности человека, которые малозаметны или даже не осознаются при восприятии исходного изображения.

Выводы

Несмотря на относительную простоту вейвлетной пространственно-частотной фильтрации для предварительной обработки изображений удалось изучить деятельность “зрительного мозга” с учетом свойства рецептивных полей нейронов зрительной коры и в условиях эксперимента проследить особенности выполнения исследуемых задач.

Так, в зависимости от семантического признака “Знакомое”/“Незнакомое” при восприятии лиц конструкция “зрительного мозга” использует разные пространственно-частотных характеристики.

Экономичное распределение “информационно-вычислительных” ресурсов головного мозга позволяет успешно осуществлять выделение значимых для наблюдателя свойств воспринимаемой сцены “Знакомое лицо” за счет горизонтальных составляющих, однако при восприятии сцены “Незнакомое лицо” значительная роль отводится диагональным составляющим.

Особенности работы фильтров выявляются и при сравнении восприятия социального фактора “Возраст”, идентифицируемого при помощи диагональных фильтров, с эмоциональным фактором “стресс”, который может быть определен уже по минимальным признакам, идущим от горизонтальных составляющих.

В результате длительной эволюции зрительная система оказалась оптимально приспособленной к восприятию пространственно-частотных свойств окружающей среды, которые не только определяют формы и положение объектов в поле зрения человека, но и выделяют их эмоциональные и социальные признаки.

Исследование выполнено за счет гранта Российского научного фонда (проект № 14-50-00069) Санкт-Петербургскому государственному университету.

* * * * *

ЛИТЕРАТУРА

1. Dakin S.C., Watt R.J. Biological “bar codes” in human faces // J. Vision. 2009. V. 9(4):2. P. 1–10. <http://journalofvision.org/9/4/2/>, doi:10.1167/9.4.2.
2. Логунова Е.В., Шелепин Ю.Е., Пронин С.В. Моделирование работы пространственно-частотных фильтров при восприятии сложных динамических сцен // Оптический журнал. 2014. Т. 81. № 11. С. 62–68.
3. Weiman C. Efficient discrete Gabor functions for robot vision // SPIE Conference on Wavelet Applications. Orlando, FLA, April, 1994. V. 2242. P. 148–160.
4. Филд Д. Согласованные фильтры, вейвлеты и статистика натуральных сцен // Оптический журнал. 1999. Т. 66. № 9. С. 25–36.
5. Жеребко А.К., Луцив В.Р. Согласованная фильтрация в естественных и искусственных нейронных сетях // Оптический журнал. 1999. Т. 66. № 9. С. 69–72.
6. Мурыгин К.В. Оптимизация габоровских вейвлет-преобразований для задачи распознавания человека по изображению лица // Искусственный интеллект. 2003. № 4. С. 223–229.
7. Viola P., Jones M. Robust real-time face detection // Int. J. Computer Vision. 2004. V. 57(2). P. 137–154.
8. Шелепин Ю.Е., Фокин В.А., Меньшикова С.В., Борачук О.В., Коскин С.А., Соколов А.В., Пронин С.В., Харузов А.К., Васильев П.П., Вахрамеева О.А. Методы иконики и методы картирования мозга в оценке функционального состояния зрительной системы // Сенсорные системы. 2014. Т. 28. № 2. С. 61–75.
9. Сборник фотографий. Электронный доступ: <http://photoshtab.ru/2011/12/felicie/>.
10. Красильников Н.Н., Шелепин Ю.Е., Красильникова О.И. Применение принципов оптимального наблюдателя при моделировании зрительной системы человека // Оптический журнал. 1999. Т. 66. № 9. С. 17–24.
11. Мирошников М.М. Согласованная фильтрация при зрительном восприятии и информационное согласование в иконике // Оптический журнал. 1999. Т. 66. № 9. Р. 5–8.