

КОГНИТИВНАЯ НАУКА В МОСКВЕ  
**НОВЫЕ ИССЛЕДОВАНИЯ**



**МАТЕРИАЛЫ  
КОНФЕРЕНЦИИ  
2017**

ПОД РЕД. Е.В. ПЕЧЕНКОВОЙ, М.В. ФАЛИКМАН

УДК 159.9

ББК 81.002

К57

К57 Коллективный

Когнитивная наука в Москве: новые исследования. Материалы конференции 15 июня 2017 г.

Под ред. Е.В. Печенковой, М.В. Фаликман. – М.: ООО «Буки Веди», ИППИП. 2017 г. – 596 стр.

Электронная версия

ISBN 978-5-4465-1509-7

УДК 159.9

ББК 81.002

ISBN 978-5-4465-1509-7

© Авторы статей, 2017

## SCAN-PATH VS. ЗРИТЕЛЬНЫЕ АТТРАКТОРЫ: ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА ГИПОТЕЗЫ НОТТОНА И СТАРКА

О. В. Левашов

[olevashov@gmail.com](mailto:olevashov@gmail.com)

ФГНБУ Научный центр неврологии

**Аннотация.** Предложена методика регистрации траектории смещения внимания на ранней стадии зрительного восприятия (preattentive vision). Пятидесяти двум испытуемым в случайном порядке трижды предъявляли набор из пяти изображений. Первая экспозиция каждого изображения составляла 300 мс, при каждом последующем предъявлении экспозиция увеличивалась на 300 мс. Таким образом по отчетам испытуемых (по постэкспозиционной матрице) можно было оценить местоположение первых трех локусов внимания. Гипотеза Ноттона и Старка о жестких траекториях осмотра (scan-path) не подтвердилась. Полученные данные говорят о наличии на изображениях «зрительных аттракторов» — зон, на которых останавливают внимание все испытуемые. Наиболее значимыми аттракторами оказались лица и фигуры людей, фигуры животных, отражения в воде и искаженные по форме изображения объектов. Эти результаты позволяют предположить, что выбор первых точек фиксации меньше связан с визуальным контекстом» изображения сцены, а больше зависит от взаимодействия между собой на высших корковых уровнях сложных нейронных комплексов («следов» памяти), сформированных в процессе обучения. Это заставляет отказаться от классической схемы зрительного узнавания на базе выделения сначала простых, а потом все более сложных локальных признаков формы, последующего их «сшивания» и сравнения с «эталоном».

**Ключевые слова:** зрительное узнавание, выбор первых точек фиксации, зрительные аттракторы, зрительное внимание, траектории осмотра изображения, модели узнавания

В зрительном восприятии часто выделяют «раннюю» (preattentive) стадию, когда движения глаз выполняются произвольно (Malik, Perona, 1990; VanRullen, Thorpe, 2001). Осознание наступает позднее, и тогда движения глаз выполняются в соответствии с решаемой задачей.

Вопрос о том, на какие точки сцены направляются первые саккадические ДГ, до сих пор остается без ясного ответа (Henderson, 2003). Ноттон и Старк (Ноттон, Старк, 1974) попытались «развернуть» процесс узнавания во времени, используя стимулы в виде крупных и неярких изображений и обнаружили, что для конкретного человека траектория первых 2–3 скачков повторяется каждый раз при показе одного и того же изображения. Для другого изображения эта траектория иная, но тоже характерная. Такие траектории они назвали «пу-

тиями сканирования» (scan-path) и предположили, что на изображениях есть ключевые «локальные признаки», которые зрительный мозг в процессе обучения распознает и кодирует, включая в описание и команды для движений глаз по нужной «траектории сканирования». Эти признаки и команды для движений глаз группируются в «кольца признаков», которые затем используются в качестве «эталонов» при узнавании.

Другой подход к проблеме выбора «интересных» точек на изображениях заключается в том, чтобы построить так называемую «функцию информативности» изображения, в максимумах или минимумах которой глаз останавливается и происходит считывание информации (Завалишин, Мучник, 1974). Для простых линейных изображений типа цифр (но не для цветных сложных) авторы смогли математически описать функцию информативности и получить экспериментальные подтверждения своей модели.

В данной работе описаны результаты проверки гипотезы Ноттона и Старка на сложных изображениях (репродукции картин, реклама, фото реальных сцен). Гипотеза Ноттона и Старка не подтвердилась, однако было обнаружено, что зрительная система располагает хранящейся в памяти своеобразной «коллекцией эталонов форм», которые настроены на выделение зрительных «аттракторов», на которые в первую очередь направляется внимание. Наиболее значимыми зрительными аттракторами оказались лица и фигуры людей, животных, сложные фрагменты знакомых предметов, а также искаженные и необычные формы. Предварительные результаты экспериментов были представлены на ECVF-2006 (Levashov, Rumyantseva, 2006).

## Методика

**Испытуемые и стимулы.** Участниками экспериментов были 52 человека, с нормальным или скорректированным до нормы зрением в возрасте от 19 до 75 лет. Использовали 9 тематических наборов изображений, включавших цветные фото реальных сцен, интерьера, малоизвестные картины известных художников и визуальную рекламу. Стимулы имели размер 20–24°. Фиксационный крестик экспонировался на 1 сек в точке, совпадающей с центром тестового изображения, которое появлялось через 100 мс после исчезновения крестика.

**Аппаратура.** Стимулы предъявлялись на экране монитора с помощью компьютерной программы, позволяющей варьировать длительность экспозиции и порядок предъявления стимулов в серии (Levashov, Rumyantseva, 2006; Левашов, 2015).

**Процедура.** Испытуемый сидел перед экраном, на котором в случайном порядке по очереди экспонировались пять изображений из одного набора. Каждый раз после показа стимула в том же месте экрана на 500 мс подавалась «постэкспозиционная матрица» такого же формата, как и стимул. Матрица имела 9 одинаковых зон разного цвета, внутри которых были размещены очень мелкие цифры от 1 до 9. Сессия для каждого испытуемого состояла из трех серий, в каждой из которых в случайном порядке предъявлялись

одни и те же пять изображений из одного набора (например, картины С. Дали). В тренировочной сессии испытуемым давали тестовый набор из пяти изображений и подбирали начальное время экспозиции — 300 либо 400 мс. Эта экспозиция была выбрана исходя из величины латентного периода скачка глаза (200–250 мс) и длительности фиксации глаза (200–300 мс) (Ярбус, 1967; Белопольский, 2005; Findlay, Gilchrist, 2005). Это гарантировало выполнение одной саккады и последующей фиксации стимула. В первой серии каждое изображение предъявлялось на 300 или 400 мс, во второй серии время экспозиции удваивалось, в третьей — утраивалось. Это позволяло испытуемому совершить одну, две или три фиксации, а экспериментатору — отследить траекторию смещения внимания по изображению. Точность определения «фокусов внимания» составляла 1.1–1.3° (половина расстояния между соседними цифрами в постэкспозиционной матрице), что близко к размеру зоны максимальной остроты зрения человека.

**Задача.** Испытуемый должен был набрать на клавиатуре цвет и цифру, которые он успевал различить. В силу малого размера цифр испытуемый видел только те цифры, которые попадали в область его фoveального зрения. Таким образом, определялась локализация первых трех локусов внимания и, соответственно, траектория смещения зрительного внимания.

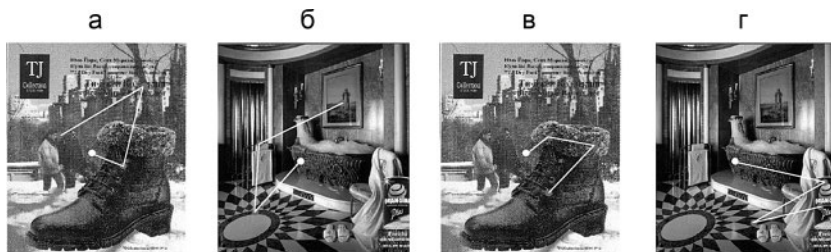
## Результаты и их обсуждение

**Пространственная структура сцены.** Многие испытуемые при показе изображений сцен не смогли дать даже примерное описание типа сцены. Например, не могли сказать, что был показан берег моря (картина С. Дали «Постоянство памяти») даже при экспозициях 800 мс. Это означает, что зрительный анализ сложных сцен требует многих скачков и фиксаций глаз.

**Зрительные аттракторы.** Оказалось, что на этом «бессознательном» этапе анализа сцены внимание практически всех испытуемых останавливается на одних и тех же характерных деталях. Мы обозначили эти фрагменты термином «зрительные аттракторы». При этом сами аттракторы оказались неравноценными. Количественный анализ выбранных в процессе узнавания зрительных аттракторов позволил составить их примерный рейтинг в порядке убывания «значимости»:

1. Искаженные предметы и необычные объекты, напоминающие по форме знакомые объекты (например, некоторые предметы на картинах С. Дали), отражения предметов в воде или в других средах.
2. Лица людей, фигуры человека или животных, части тела, знакомые предметы.
3. Крупные надписи и крупные контрастные текстуры (например, решетки).

Контрастные одиночные объекты (например, камни в воде) также могли привлечь внимание испытуемых, но только при отсутствии более важных зрительных аттракторов.



**Рисунок 1.** Типичные траектории внимания у «визуала» (а и б) и у «аналитика» (в и г). Заметно, что «визуал» улавливает «структуру» сцены, фиксируя передний, средний и задний план, тогда как внимание «аналитика» ограничено частью сцены.

**Scan-path.** Траектории движения внимания по изображениям были разными, даже при повторном показе того же изображения одному испытуемому. Таким образом, гипотеза о повторяющихся траекториях осмотра («scan-path») не подтвердилась.

**«Визуалы» и «аналитики».** По характеру пространственного распределения зрительного внимания всех испытуемых можно было разбить на две группы – условно назовем их «аналитики» и «визуалы». «Визуалы» могли оценивать тип сцены (рис. 1 а, б), они замечали больше зрительных аттракторов, и траектории движения их внимания были достоверно более протяженными, чем у «аналитиков», охватывая гораздо большее «поле внимания», вплоть до всего изображения. В отличие от «визуалов», «аналитики» не могли оценивать пространственную конфигурацию сцены, их внимание было ограничено частью изображения (рис. 1 в, г). Подробные количественные данные, описывающие различия «визуалов» и «аналитиков» приведены в нашей недавней работе (Левашов, 2017).

Полученные данные показывают, что за время порядка всего 50–100 мс после появления изображения зрительная система способна оценить информативность большей части сцены и найти там наиболее важные объекты (зрительные аттракторы). Эта цифра получается из сравнения минимальной длительности экспозиции (300 мс) и величины латентного, необходимого для программирования кинематики скачка (200–250 мс). Это может означать наличие параллельного анализа изображения сцены с помощью набора «нейронных эталонов» в зрительной памяти. Маловероятно, что за такое короткое время можно выполнить последовательный поиск и выделение «локальных признаков», построение множества «описаний» и сравнение этих описаний с «эталонами».

### **Зрительные аттракторы и моделирование узнавания**

В настоящее время в моделировании процесса зрительного анализа сцен можно выделить три подхода. Первый основан на оценке «визуального контекста», то есть наличия на изображениях специфических локальных признаков, которые зрительная система обнаруживает, собирает в некоторое «описание»

и далее сравнивает с «эталоном» в памяти (Ноттон, Старк, 1974). Второй подход предполагает сканирование изображения сцены под управлением «эталона» памяти в рамках конкретной задачи, решаемой человеком. Этот подход получил ряд экспериментальных подтверждений. Так, Стейн и др. (Stain et al., 2017) показали, что узнавание гораздо сильнее зависит от сформированных в процессе опыта «эталонов», чем от выделения локальных признаков форм. Третий подход – комбинация первых двух, когда один блок выделяет локальные признаки, а другой по набору этих признаков выбирает в памяти наиболее близкий «эталон», который задает дальнейший анализ сцены (Левашов, 1989). Модель узнавания, учитывающая такую комбинацию и параллельный анализ сцены, описана в недавней работе автора (Левашов, 2017).

## Литература

- Белопольский В. И. Взор человека. М.: ИПРАН, 2007.
- Завалишин Н. В., Мучник И. Б. Модели зрительного восприятия и алгоритмы анализа изображений. М.: Наука, 1974.
- Левашов О. В. Вычислительные модели сенсорных систем. М.: ВИНТИ, 1989.
- Левашов О. В. Способ определения траектории смещения произвольного зрительного внимания на изображении. Патент на изобретение РФ № 2545428 от 24.02.2015.
- Левашов О. В. Атракторы внимания на ранней стадии зрительного анализа сцен и новая модель узнавания // Вестник государственного университета «Дубна». 2017. № 1. URL: <http://vestnik.fsgn.uni-dubna.ru/issues/Levashov.pdf>.
- Ноттон Д., Старк Л. Движения глаз и зрительное восприятие // Восприятие. Механизмы и модели. М., 1974. С. 226–240.
- Ярбус А. Л. Роль движений глаз в процессе зрения. М.: Наука, 1965.
- Findlay J. M., Gilchrist I. D. Active vision: The psychology of looking and seeing. Oxford University Press, 2005.
- Henderson J. M. Human gaze control during real-world scene perception // Trends in Cognitive Sciences. 2003. Vol. 7. No. 11. P. 498–504. doi:10.1016/j.tics.2003.09.006
- Levashov O., Romyantseva R. Visual attractors for eye movements in early vision // Perception. 2006. Vol. 35. P. 86.
- Malik J., Perona P. Preattentive texture discrimination with early vision mechanisms // Journal of the Optical Society of America A. 1990. Vol. 7. No. 5. P. 923–932. doi:10.1364/josaa.7.000923
- Stain T., Peelen M. V. Object detection in natural scenes: Independent effects of spatial and category-based attention // Attention, Perception, & Psychophysics. 2017. Vol. 79. No. 3. P. 738–752. doi:10.3758/s13414-017-1279-8
- VanRullen R., Thorpe S. The time course of visual processing: From early perception to decision-making // Journal of Cognitive Neuroscience. 2001. Vol. 13. No. 4. P. 454–461. doi:10.1162/08989290152001880

## Scan-Path vs Visual Attractors: Experimental Evaluation of the Notton-Stark Hypothesis

Levashov O.V.

[olevashov@gmail.com](mailto:olevashov@gmail.com)

Research Center of Neurology, Moscow, Russia

**Abstract.** The results of experimental verification of the hypothesis of Notton and Stark for complicated images of spatial scenes (e.g. paintings reproductions, advertising) are described. However, we found that human visual attention in early (pre-attentive) vision is based on a hierarchy of visual “attractors” (VAT), where attention and consequently eye movements are directed first. The strongest VATs turned out to be areas of images that include complicated or unusual forms, such as reflections in water or distorted real objects (as in the paintings of the Surrealists). The trajectories of visual attention of the image pass through the different VATs depending on their perceptual “weights” and therefore can vary.

**Keywords:** visual attractors, early (pre-attentive) vision, analysis of visual scenes, scan-path, models of vision