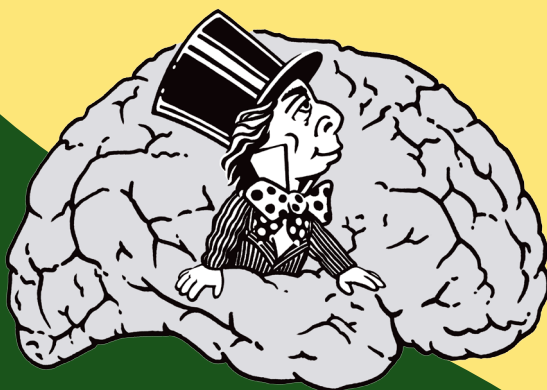


КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ
КОНФЕРЕНЦИИ
2023

Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 21 – 22 июня 2023 г. Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман. – М.: ООО «Буки Веди», Московский институт психоанализа. 2023 г. – 604 стр.

© Авторы статей, 2023

ISBN 978-5-4465-3880-5

УДК 159.9
ББК 88.25

ISBN 978-5-4465-3880-5

© Авторы статей, 2023

РОЛЬ ОБЛАСТЕЙ ЗРИТЕЛЬНОЙ СЦЕНЫ С НАИБОЛЬШИМ ПРИРОСТОМ НЕЛОКАЛЬНОГО КОНТРАСТА В ЭКЗОГЕННОМ УПРАВЛЕНИИ ВНИМАНИЕМ

Д. В. Явна*, В. В. Бабенко, Е. Г. Родионов
yavna@fortran.su
ЮФУ, Ростов-на-Дону

Аннотация. Неравномерность распределения фиксаций взгляда на изображении была наглядно продемонстрирована уже в работах А. Л. Ярбуса с соавторами. Известно, что начальные фиксации направляются самим изображением, при этом «точками интереса» являются области, отличающиеся по своим физическим характеристикам от окружения. Мы предположили, что участками изображений, которые притягивают внимание наблюдателя, могут быть области с наибольшим приростом контраста относительно окружения. С целью проверки этой гипотезы регистрировались движения глаз испытуемых при рассматривании изображений сцен в ходе решения задачи определения их эмоциональной валентности. Затем, с помощью разработанной программы, в стимульных изображениях на разных пространственных частотах были найдены области с наибольшим, наименьшим и промежуточным приростом суммарного контраста. Сопоставление карт распределения фиксаций с картами распределения найденных областей показало, что области, на которых фиксировался взгляд, в существенной степени совпадают с областями сцен, характеризующимися наибольшим приростом контраста.

Ключевые слова: зрительное внимание, визуальная салиентность, движения глаз, айтрекинг, моделирование внимания, области интереса

Исследование поддержано Российским научным фондом (проект № 20-64-47057).

Введение

Движения глаз часто сопровождают смещение внимания в новую область поля зрения (Theeuwes, 2014). Поскольку выбор цели внимания основан на принципе максимизации информации (Bruce, Tsotsos, 2005), распределение фиксаций взгляда при рассматривании изображений указывает на области, которые являются наиболее информативными. Существует точка зрения, что областями интереса в изображении являются области, отличающиеся по своим физическим характеристикам от окружения (Fuchs et al., 2011; Gao et al., 2009; Hou et al., 2007). Особенностью нашего подхода является то, что мы рассматриваем в качестве областей интереса участки изображения с наибольшим приростом контраста по сравнению с окружающей областью, которые могут быть обнаружены зрительными механизмами второго порядка (Graham, 2011) на разных пространственных частотах; сами же механизмы могут быть кандидатами на роль «ворот внимания» (Бабенко, 2023).

Целью исследования является изучение связи между локализацией фиксаций взгляда при рассматривании сцен в задаче определения их эмоциональной валентности и пространственным распределением в изображении

областей с наибольшим приростом нелокального (суммарного по участку изображения) контраста.

Методика

В экспериментах приняли участие 100 студентов университета в возрасте 19–21 года с нормальным зрением. В качестве стимулов использовались 360 фотографий зрительных сцен. Стимулы представляли собой коллекцию фотографий сцен, предварительно оцененных экспертами как имеющие негативную, нейтральную или позитивную эмоциональную окраску. Испытуемые должны были дать собственную оценку нажатием одной из трех клавиш. Эти результаты в данном исследовании не учитывались, а сама постановка задачи имела целью организовать работу испытуемого. При рассмотрении изображений, предъявляемых на 700 мс, с помощью айтрекера SMI Red-m регистрировались движения глаз наблюдателей. Время регистрации было ограничено для того, чтобы минимизировать влияние эндогенных факторов на распределение фиксации. По объединенным результатам всех испытуемых строились бинарные карты фиксации взора и тепловые карты плотности фиксации для каждого стимульного изображения. Затем, с помощью разработанной программы, в стимульных изображениях были найдены области с наибольшим (max), наименьшим (min) и промежуточным (med) приростом суммарного нелокального контраста по сравнению с окружением. Оценка прироста контраста проводилась с помощью концентрического скользящего окна, разделенного на центр и периферию, представляющего собой модель зрительного механизма второго порядка. Найденные области были представлены на карте двумерными гауссианами. Карты распределения указанных областей строились для пяти диапазонов пространственных частот с центральной частотой 4, 8, 16, 32 и 64 цикла на изображение, а затем объединялись в единую вероятностную карту, причем размер гауссиана был жестко связан с пространственной частотой. Пример изображения и построенных для него карт показан на рис. 1.

Модельные вероятностные карты сравнивались с эмпирическими (бинарными картами фиксации и тепловыми картами) при помощи следующих метрик:

- Площадь под кривой РХП (AUC, area under ROC curve) в варианте, предложенном Т. Джадд с соавторами (Judd et al., 2009). Данная метрика используется для сравнения модельных вероятностных карт с бинарными картами фиксации. Она сильно «штрафует» за ложноположительные прогнозы, предполагающие большую вероятность фиксации, и в малой степени реагирует на слабые ложноположительные прогнозы (Bylinskii et al., 2019).
- Нормализованная салиентность пути осмотра (NSS, normalized scanpath saliency), предложенная Р. Дж. Питерсом с соавт. (Peters et al., 2005). Как и AUC, применяется для сравнения модельных вероятностных карт с бинарными картами фиксации. Она в примерно равной мере «штрафует» как за ложноположительные, так и за ложноотрицательные прогнозы,

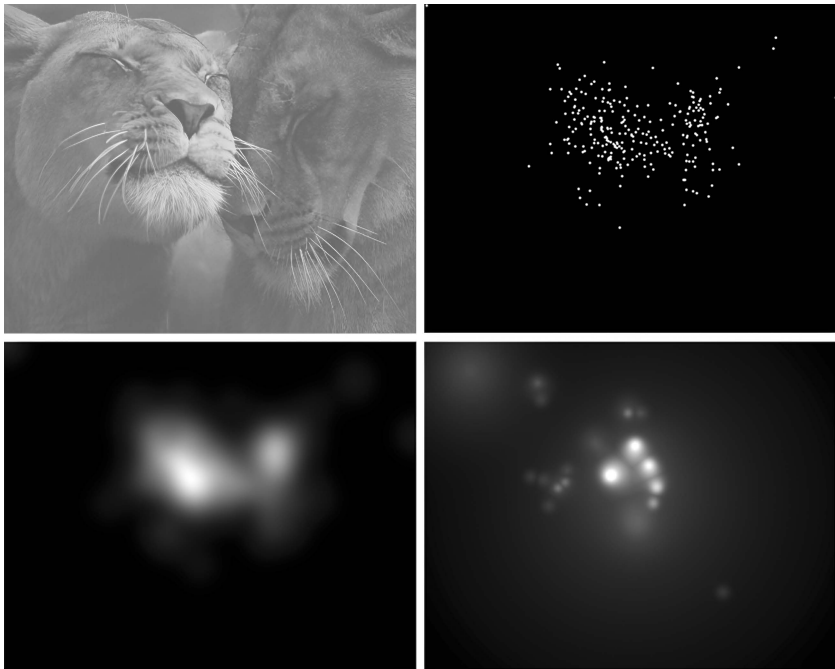


Рисунок 1. Пример стимульного изображения и соответствующих ему карт. Слева направо, сверху вниз: стимульное изображение; бинарная карта фиксаций; вероятностная тепловая карта плотности фиксаций; модельная вероятностная карта

- будучи, в отличие от AUC, хорошо чувствительной и к слабым ложноположительным предсказаниям (Bylinskii et al., 2019).
- Коэффициент корреляции Пирсона (CC). Как и NSS, CC является хорошо сбалансированной в отношении ложноположительных и ложноотрицательных прогнозов метрикой (Bylinskii et al., 2019). В отличие от AUC и NSS, CC используется для сравнения модельных вероятностных карт не с бинарными картами фиксаций, а с тепловыми картамию.
 - Дивергенция Кульбака-Лейблера (KL) — мера, используемая в теории информации для оценки различий двух вероятностных распределений. Применяется для тепловых карт и сильно «штрафует» за ложноотрицательные прогнозы (Bylinskii et al., 2019). Это именно мера *дивергенции*, и чем выше ее значение, тем больше различий между модельной картой saliентности и картой, построенной по фиксациям.

Значения каждой из четырех метрик, вычисленные для max, med и min, сравнивались между собой с использованием критерия Краскела-Уоллиса с целью проверки гипотезы о сдвиге центральной тенденции. Каждая из срав-

ниваемых между собой трех независимых «групп» включала по 360 значений конкретной метрики (по числу сравниваемых пар карт).

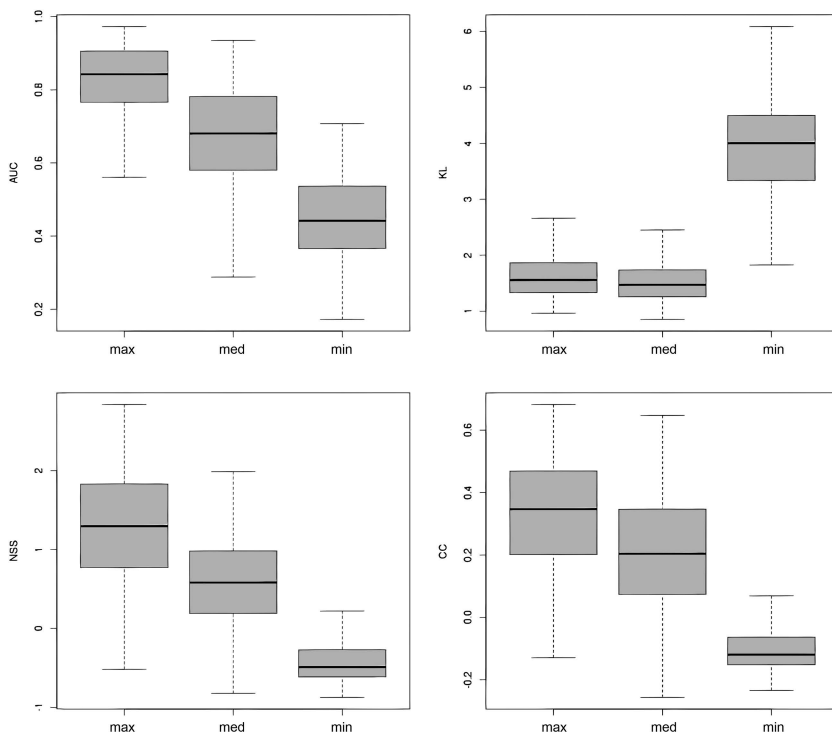


Рисунок 2. Разброс значений метрик для модельных карт max, med и min. «Ящик» ограничивает квартили 2 и 3, концы «усов» соответствуют границам диапазона наблюдаемых значений, жирная черта – медиане

Результаты

Медианные значения AUC, NSS и CC значительно снижались в последовательности max > med > min. Для AUC медианы составили соответственно 0.845, 0.681 и 0.442 ($H(2) = 688.902$, $p < .001$, для NSS – 1.296, 0.582 и -0.491 ($H(2) = 724.271$, $p < .001$), для CC – 0.347, 0.204 и -0.12 ($H(2) = 656.886$, $p < .001$). Для KL эта последовательность нарушалась: дивергенция между тепловыми картами и модельными картами med была меньше (1.472), чем между тепловыми картами и max (1.557), но для min она тем не менее была велика (4.003). Эта разница показателей была значимой ($H(2) = 710.091$, $p < .001$), причем постхок-тест Данна обнаружил значимые различия и между max и med ($p = .009$). Полученные результаты показаны на рис. 2.

Обсуждение и выводы

Полученные результаты указывают на то, что области фиксации взора при рассмотривании сцен значительно совпадают с расположением областей с наибольшим приростом нелокального контраста. По мере уменьшения амплитуды модуляции контраста выделенных областей в ряду *max-med-min* эта связь значительно снижается. Отклонение метрики Кульбака-Лейблера от этой последовательности снижения может объясняться ее повышенной чувствительностью к ложноотрицательным прогнозам.

Таким образом, обнаружено, что области фиксации взора при определении эмоциональной валентности сцены в существенной степени совпадают с областями изображения, характеризующимися наибольшим приростом нелокального контраста. Полученные результаты свидетельствуют о том, что целыми внимания при восприятии зрительных сцен могут быть области с наибольшим приростом суммарного контраста.

Литература

- Бабенко В.В.* Зрительные механизмы второго порядка: обзор исследований // Журнал высшей нервной деятельности им. И. П. Павлова. 2023. Т. 73. № 1. С. 3 – 23.
- Bruce N., Tsotsos J.* Saliency based on information maximization // NIPS'05: Proceedings of the 18th International Conference on Neural Information Processing Systems. 2005. P. 155 – 162.
- Bylinskii Z., Judd T., Oliva A., Torralba A., Durand F.* What do different evaluation metrics tell us about saliency models? // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. Vol. 41. No. 3. P. 740 – 757. <https://doi.org/10.1109/tpami.2018.2815601>
- Fuchs I., Ansorge U., Redies C., Leder H.* Saliency in paintings: Bottom-up influences on eye fixations // Cognitive Computation. 2011. Vol. 3. No. 1. P. 25 – 36. <https://doi.org/10.1007/s12559-010-9062-3>
- Gao D., Han S., Vasconcelos N.* Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. Vol. 31. No. 6. P. 989 – 1005. <https://doi.org/10.1109/tpami.2009.27>
- Graham N.V.* Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years // Vision Research. 2011. Vol. 51. No. 13. P. 1397 – 1430. <https://doi.org/10.1016/j.visres.2011.02.007>
- Hou X., Zhang L.* Saliency detection: A spectral residual approach // 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA: IEEE, 2007. P. 1 – 8. <https://doi.org/10.1109/cvpr.2007.383267>
- Judd T., Ehinger K., Durand F., Torralba A.* Learning to predict where humans look // 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009. P. 2106 – 2113. <https://doi.org/10.1109/iccv.2009.5459462>
- Peters R.J., Iyer A., Itti L., Koch C.* Components of bottom-up gaze allocation in natural images // Vision Research. 2005. Vol. 45. No. 18. P. 2397 – 2416. <https://doi.org/10.1016/j.visres.2005.03.019>
- Theeuwes J.* Spatial orienting and attentional capture / A.C. (Kia) Nobre, S. Kastner (Eds.). Oxford, UK: Oxford University Press, 2014. <https://doi.org/10.1093/oxfordhb/9780199675111.013.005>

THE ROLE OF IMAGE AREAS WITH THE HIGHEST INCREASE IN NON-LOCAL CONTRAST IN EXOGENOUS ATTENTION CONTROL

D. V. Yavna*, V. V. Babenko, E. G. Rodionov

yavna@fortran.su

Southern Federal University, Rostov-on-Don

Abstract. The inhomogeneity of gaze fixation distribution on an image has been clearly demonstrated in the works of A. L. Yarbus et al. It is known that the initial fixations are guided by the image itself, while “points of interest” are areas that differ in physical characteristics from their surroundings. We assumed that the areas of images that attract an observer’s attention may be those with the greatest increase in total luminance contrast relative to the surrounding areas. In order to test this hypothesis, eye movements were recorded while an observer viewed images of scenes in the task of determining their emotional valence. Then, using our program, those areas with the maximum, minimum, and intermediate increases in total luminance contrast were located in stimulus images at different spatial frequencies. A comparison of fixation distribution maps with distribution maps of the identified areas showed that the areas of gaze fixation largely coincided with the areas of the scenes characterized by the maximum increase in total luminance contrast.

Keywords: visual attention, visual saliency, eye movements, eye tracking, attention modeling, areas of interest

Research supported by the Russian Science Foundation (Project No. 20-64-47057).