ЭМПИРИЧЕСКИЕ ИССЛЕДОВАНИЯ

Структура реалистичных сцен как фактор, направляющий зрительный поиск

Кукарцев Д.С.[1], **Морозов М.И.**[1]

[1] Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации, Москва, Россия

Зрительный поиск в сценах направляется рядом факторов, в основе которых лежат разные типы информации о них. Одним из них выступает зрительная заметность, яркость, которая опирается на зрительные признаки сцены и направляет внимание от наиболее заметной области к наименее заметной. Другим фактором выступает смысл и структура сцены. В нем рассматривается согласованность объекта и сцены, в которой он находится. Согласованность может выступать как высоковероятная принадлежность объекта сцене (смысл) и как пространственно ожидаемое расположение объектов (структура). Данная работа будет посвящена обзору этих факторов, ограничениям ранних моделей зрительного поиска в сценах, сопоставлению модели Дж. Вольфа с результатами недавних исследований структурно-смыслового фактора и фиксации ряда возникших теоретических вопросов.

Ключевые слова: зрительный поиск, структура сцены, смысл сцены, иерархия сцены, якорные объекты, перцептивная заметность

Введение

Зрительный поиск — это класс задач, используемых в исследованиях зрительного внимания. Испытуемым требуется найти целевой объект среди множества отвлекающих стимулов — дистракторов [Фаликман, 2006]. Для этого необходимо произвести отбор релевантной информации, для чего как раз и необходимо внимание. Центральной проблемой в исследованиях зрительного поиска является вопрос о том, каким образом происходит отбор релевантной информации и какие факторы на него влияют. В классических работах исследователи были сконцентрированы на изучении влияния признаков объектов (цвета, формы, наклона и т. д.) на эффективность зрительного поиска. Механизмы этого влияния наиболее полно описываются в теории управляемого поиска [Wolfe, 2021]. В рамках нее зрительный поиск представляется как двухэтапный процесс. На первом этапе происходит параллельный анализ всех зрительных признаков предъявленных объектов. На втором этапе внимание последовательно направляется на те места в пространстве, в которых находятся признаки, релевантные цели поиска. Например, если целью поиска является красный квадрат, предъявленный среди красных и черных кругов и квадратов, то внимание будет направляться на все красные объекты и все квадраты. Механизм отбора релевантной информации при помощи параллельного анализа зрительных признаков получил название гайденса (guidance). Теоретическая ценность этой идеи заключается в описании того, каким образом будет происходить отбор информации еще до того, как на нее будет обращено внимание, возможности которого ограничены.

Описанные выше механизмы исследовались на материале задач, в которых предъявленные объекты были никак не связаны между собой. Однако зрительный поиск в реальной жизни протекает совсем в других условиях. Мы ищем объекты, находящиеся в определенных локациях (кухня, ванная, офис, улица и т. п.), называемых сценами (scene), а не на одноцветном фоне, как в лабораторных задачах. В сценах объекты располагаются не хаотично, а в соответствие со структурными и смысловыми закономерностями. Например, кружки стоят на столах, а не парят в воздухе, и вряд ли встречаются в ванной. Существуют данные, демонстрирующие, что зрительный поиск в сцене оказывается эффективнее, чем среди тех же объектов, расположенных в случайном порядке [Wolfe et al., 2011]. Но эти данные представляли проблему для ранних версий теории управляемого поиска [Wolfe, 1994], потому что на уровне параллельной обработки зрительных признаков не важно, к какому объекту эти признаки относятся и в какой сцене находятся. Считалось, что для распознавания объектов и самой сцены необходимо внимание, а значит, их свойства не могли направлять внимание в процессе поиска. Для решения этой проблемы исследователи предлагают различные факторы, вносящие свой вклад в направление внимания в сценах, а также механизмы их работы. Целью нашего обзора является рассмотрение и анализ влияющей на зрительный поиск информации о сцене, а именно: освещение теоретических положений ранних моделей и их ограничений, послуживших поводом для изучения и введения новых направляющих внимание свойств сцены; описание этих новых свойств — карты значений, контекста как глобальной структуры сцены и контекста как совместной встречаемости объектов в сцене; соотнесение эмпирических результатов исследований с моделью управляемого поиска Дж. Вольфа, а также фиксация проблемных вопросов, на которые только предстоит ответить.

Коротко предваряя последующий обзор, сформулируем имеющиеся на данный момент позиции и отметим, что среди исследователей нет согласия относительно того, что оказывает доминирующее влияние на направление внимания в сценах [Chan, Hayward, 2013; Võ, Boettcher, Draschkow, 2019]. Одна группа теорий говорит о том, что ключевую роль играет зрительная заметность объектов [Itti, Koch, 2001]. Согласно модели Л. Итти и Х. Коха в результате обработки сцены формируется карта заметности (saliency map) — топографическое распределение перцептивной яркости, заметности участков сцены. Внимание при этом перемещается от наиболее заметного к наименее заметному участку [Itti, Koch, 2001].

Другая группа теорий отводит главную роль смыслу и структуре сцены, которые в ряде работ метафорически называются «семантикой» и «синтаксисом» [Võ, 2021; Võ, Boettcher, Draschkow, 2019]. Идея заключается в том, что у нас есть набор правил, обладающих общими чертами с языковыми структурами [Измайлов, Черноризов, 2005], которые позволяют эффективно осуществлять поиск. Здесь же возникает идея о карте значений (meaning map) — количественном распределении осмысленности в изображении [Henderson, Hayes, 2017]. Так, например, фрагменты изображения, где представлена однородная поверхность объекта (например, шторы), оцениваются как низкие по осмысленности, в то время как фрагменты изобра-

жения с целостными объектами (тарелка) — как высокие [Henderson, Hayes, 2017].

Зрительная заметность как фактор, направляющий внимание: модель Итти и Коха

Возможности нашего внимания ограничены, и мы не способны видеть абсолютно все в сцене ясно и отчетливо. Для этого нам приходится направлять его в определенные участки сцены, чтобы проверить, находится ли в них нужный нам объект. При этом было бы неэффективно осуществлять поиск во всех возможных местах сцены, т. к. это сопровождалось бы повышенными временными затратами. Что же в таком случае делает поиск эффективным? Один из ответов дает основанная на работе зрительной системы модель Итти и Коха [Itti, Koch, 2001]. Она предлагает возможный механизм работы восходящего (bottom-up) внимания, т. е. такого, которое направляется зрительными признаками сцены. Как разворачивается этот процесс? Во-первых, зрительная система параллельно обрабатывает всю сцену по отдельным признакам (цвет, ориентация, интенсивность и т. д.). Во-вторых, каждый из каналов признаков рассчитывает локальные контрасты, т. е. области, обладающие повышенной перцептивной заметностью за счет более сильного нейронного ответа. В-третьих, результаты обработки отдельных признаков объединяются и создают карту заметности, которая задает последовательность для сканирования участков сцены вниманием. Не зацикливаться на одном участке и переключаться к другому, предположительно, позволяет механизм торможения возврата, который снижает активность нейронов, если в области, на которой сфокусировано внимание, отсутствует целевой стимул. Переключение происходит до тех пор, пока не будет найден целевой объект [Itti, Koch, 2001]. Таким образом, модель основывается на нейронных механизмах обработки перцептивных признаков и позволяет определить, как будет направляться внимание в сцене. Она действительно позволяет нам делать прогнозы относительного того, как внимание будет распределено, куда именно будет смотреть испытуемый. Однако она лучше всего предсказывает направление внимания в случаях, когда стоит задача свободного просмотра сцены, а не поиска конкретного объекта в ней. В одной из работ было показано, что при поиске объекта в реальном помещении больше фиксаций приходилось на контекстуально связанный с ним предмет, чем на объекты с наибольшей перцептивной заметностью [Mack, Eckstein, 2011]. Подобного рода данные подтолкнули исследователей к тому, что могут существовать другие, основанные на свойствах самой сцены факторы, направляющие внимание.

Смысл и структура сцены как факторы, направляющие внимание

В самом начале мы сформулировали тезис о том, что у нас есть набор правил, которые помогают находить объекты в сценах. Из него вытекает, что если правила нарушены, то обработка объектов, вероятно, должна протекать хуже, нежели в отсутствие таковых. Одна из первых работ, в которой проверялось данное предположение, принадлежит И. Бидерману [Biederman, Mezzanotte, Rabinowitz, 1982]. Ему удалось выделить типы нарушенных отношений между объектами в сцене, которые замедляют и снижают точность их распознавания [Biederman, Mezzanotte, Rabinowitz, 1982]. Для целей обзора наиболее важны следующие: 1) неправдоподобность (низкая вероятность того, что объект появится в сцене) и 2) неверное расположение (объект находится в сцене, но не в том месте, где мы ожидаем). Данные нарушения связаны с разными пониманиями контекста: в первом случае речь идет о смысловом рассогласовании объекта и сцены (например, тостер в ванной) [Davenport, Potter, 2004; Neider, Zelinsky, 2006], во втором — о контексте как пространственных отношениях между объектами (например, мышь для ноутбука располагается на экране, а не на столе) [Boettcher et al., 2018; Castelhano, Heaven, 2011; Greene, 2013; Võ, Wolfe, 2013]. Таким образом, ключевое, что можно зафиксировать из работы Бидермана: 1) локализация объектов в сцене имеет значение для их обработки и 2) между объектом и контекстом существуют разные отношения, которые могут иметь смысловую или структурную основу. В более поздних исследованиях разные трактовки контекста будут зафиксированы в явном виде [Castelhano, Heaven, 2010]. Перейдем к их рассмотрению.

Карта значений

В качестве фактора, объясняющего направление внимания в сцене, было выдвинуто и проверено предположение о наличии так называемой карты значений, смыслов [Henderson, Hayes, 2017].

Дж.М. Хендерсон и Т.Р. Хейс говорят о том, что смысл неравномерно распределен по сцене, т. е. разные ее участки воспринимаются как более или менее осмысленные, и что именно он позволяет лучше объяснить направление внимания в сценах. С помощью ряда технических процедур авторам удалось создать карты значений для разных сцен и показать, что они достаточно сильно коррелируют с картами заметности. Из этого, как отмечают авторы, следует, что эффекты, которые раньше объяснялись заметностью, могут на самом деле относиться к смыслу сцены. Дополнительно авторам удалось получить свидетельства в пользу того, что значение лучше заметности объясняет распределение фиксаций в сцене [Henderson, Hayes, 2017]. Таким образом, смысл как одна из составляющих сцены оказывает влияние на направление внимания и, вероятно, даже сильнее, чем перцептивная заметность.

Контекст как наиболее общее определение структуры сцены

В одних из первых работ, посвященных изучению влияния сцены на направление внимания, использовалось понятие контекста. Оно применялось для обозначения согласованности между целевым стимулом и сценой или ее областью. При этом согласованность могла быть смысловой или структурной, однако не во всех работах присутствует такое разделение [Neider, Zelinsky, 2006], что не позволяет понять вклад структурной составляющей сцены. В связи с этим встает задача изучить только ее влияние в отсутствие смыслового соответствия. Решение данной задачи мы находим в исследовании М.С. Кастеляно и Ч. Хейвен. Было показано, что структурная согласованность может улучшать поиск сама по себе [Castelhano, Heaven, 2011]. Объекты предъявлялись в согласованном и несогласованном контексте в согласованном и несогласованном местоположении (например, баночка для красок находилась в мастерской, но стояла не на столе, а на токарном станке). В согласованном местоположении объекты находятся быстрее и с большей точностью [Castelhano, Heaven, 2011]. Время до первой фиксации на цели и количество фиксаций, совершаемых до фиксации цели, также подтверждают эффективность поиска [Castelhano, Heaven, 2011]. Таким образом, мы можем сделать вывод, что контекст, взятый лишь в своем структурном аспекте, формирует наши ожидания относительно того, где будет располагаться объект в сцене, и направляет в соответствующие места наше внимание.

Когда мы говорим о контексте в структурном смысле, то возникает вопрос: что именно в рамках пространственного окружения направляет зрительный поиск? Один из возможных ответов — другие объекты, связанные с целевым. В работе М.Р. Грини были статистически описаны различные типы сцен с точки зрения плотности расположения, уникальности, размера, пространственной локализации, частоты встречаемости, совместной встречаемости и информативности объектов [Greene, 2013]. Данная работа показывает, что, с одной стороны, сцены неоднородны по своему наполнению, и, с другой, объекты тоже неоднородны. Они могут предоставлять разную информацию о типе сцены и о том, какие объекты могут располагаться рядом друг с другом. Так, например, в спальне совместно с кроватью чаще всего встречаются подушка, окно, лампа, потолок, штора, картина и тумбочка. Таким образом, у объектов существует разная частота встречаемости друг с другом. Это один из важных выводов, который подводит нас к направляющей функции отдельных объектов, а также к одному из пониманий контекста — совместному появлению объектов друг с другом.

Контекст как совместное появление объектов в сцене

Исследование С.С. Мака и М.П. Экштейна показывает, что совместная встречаемость объектов может влиять на зрительный поиск [Mack, Eckstein, 2011]. Испытуемым требовалось найти целевой объект (наушники), находящийся рядом с предметом, который часто встречается с целевым (MP-3 плеер). Было показано, что такие «объекты-подсказки» ускоряют поиск. Более того, когда целевой стимул находился в удаленном от объекта-подсказки месте, больше всего фиксаций было на именно на последнем. Таким образом, можно сделать вывод, что отдельные объекты способны выполнять направляющую функцию, т. е. определять приоритетные зоны для поиска.

Данная идея развивается в концепции так называемых объектов-якорей [Boettcher et al., 2018; Draschkow, Võ, 2017; Helbing, Draschkow, Võ, 2022; Turini, Võ, 2022]. Как уже было сказано ранее, объекты появляются не где угодно, а в сочетании с другими [Greene, 2013]. В более поздних работах было показано, что существуют объекты, вокруг которых группируются остальные. В исследовании М. Во и Д. Драшкова участники конструировали в виртуальной реальности сцены,

которые либо соответствуют их представлениям (согласованные), либо не соответствуют (несогласованные) [Draschkow, Võ, 2017]. Сцены составлялись из локальных и глобальных (якорных) объектов. Первые выступают как легкоподвижные и в повседневной жизни являются целями наших действий. Вторые являются статичными и, по гипотезе авторов, группируют вокруг себя локальные. Авторам удалось зафиксировать, что испытуемые чаще брали глобальные объекты первыми и сначала расставляли их при построении сцены, а только потом ставили локальные. Из этого был сделан вывод, что якорные объекты используются для создания общей структуры сцены, чтобы была возможность поместить локальные объекты в четко определенное место в ней [Draschkow, Võ, 2017].

Дополнительная проверка идеи о якорных объектах осуществляется в исследовании С.Е.П. Беттхер и коллег [Boettcher et al., 2018]. Испытуемым демонстрировались сцены с объектами-якорями и сцены с замененными якорями. На их месте располагались объекты, максимально похожие по размеру, форме и зрительной заметности.

Перед рассмотрением эмпирических результатов стоит отметить, что якоря были подобраны не интуитивно, а операционализированы через четыре метрики: 1) частота совместного появления объектов; 2) инвертированное среднее расстояние между парами объектов; 3) инвертированная пространственная вертикальная дисперсия объектов, т. е. насколько последовательно два объекта появлялись в сценах с аналогичным расположением (например, зубная щетка над раковиной); 4) частота кластеров объектов, т. е. насколько часто объекты группируются вокруг определенного якоря. Инверсия используется для того, чтобы меньшее расстояние и меньшая дисперсия были представлены большими значениями [Boettcher et al., 2018].

По результатам трех экспериментов авторам удалось установить, что благодаря якорным объектам уменьшается площадь сцены, покрываемая картой распределения внимания (scene coverage), а также повышается эффективность поиска в условиях ограниченного по времени предварительного просмотра сцены. Таким образом, можно сделать вывод, что якоря выполняют направляющую функцию и способствуют поиску целевого стимула [Boettcher et al., 2018].

Выделение локальных и глобальных объектов подводит авторов к идее об иерархической организации объектов в сцене [Võ, Boettcher, Draschkow, 2019]. Помимо указанных, еще одной структурной единицей выступают «фразы». Они представляют собой скопления локальных объектов вокруг якорей и, таким образом, любую сцену можно рассматривать как совокупность фраз. Проверка гипотезы о том, насколько данная структура соотносится с нашими реальными представлениями о сценах, будет описана ниже [Turini, Võ, 2022], пока же просто зафиксируем данные тезисы, чтобы соблюсти временную хронологию развития идей о структуре сцены.

Результаты исследования Д. Хельбинга, Драшкова и Во также свидетельствуют в пользу того, что якоря влияют на направление внимания при зрительном поиска [Helbing, Draschkow, Võ, 2022]. Авторы варьировали пространственную согласованность — последовательное либо хаотичное расположение объектов — и предъявление якоря — он мог быть заменен на серый кубоид. Участники искали объекты в виртуальной реальности. Удалось зафиксировать, что в согласованных сценах присутствие якорей облегчало поиск, в то время как в несогласованных оно нарушало управление вниманием, что отражалось в показателях движений глаз. Кроме того, в согласованных сценах якорные объекты облегчают идентификацию близлежащих целевых локальных объектов, что проявляется в скорости ответа участников. Таким образом, объекты-якоря обеспечивают эффективное направление зрительного поиска в согласованных сценах, но становятся отвлекающим фактором в несогласованных [Helbing, Draschkow, Võ, 2022].

Как было сказано ранее, ряд авторов выдвинул гипотезу о сходстве иерархии сцены с нашими представлениями, знаниями о сцене. Она была проверена в исследовании Я. Турини и Во [Turini, Võ, 2022]. Участникам предъявлялось по три изображения объектов, среди которых необходимо было выбрать, как им кажется, лишний. В результате проведения анализа репрезентативного сходства удалось установить, что объекты, принадлежащие к одной сцене, фразе или типу считались более похожими, чем объекты, принадлежащие к разным сценам, фразам или типам [Turini, Võ, 2022]. Говоря коротко, за счет принадлежности к одному иерархическому звену объекты воспринимаются как более похожие. На основе этого авторы предполагают, что данная иерархия может быть включена в семантические репрезентации объектов в памяти [Turini, Võ, 2022]. Таким образом, результаты исследования свидетельствуют в пользу того, что свойство одних объектов задавать направ-

ление поиска других может быть связано с разными репрезентациями этих объектов.

Соотнесение модели Вольфа и исследований структурно-смыслового фактора сцены, направляющего внимание

В области зрительного поиска модель Вольфа — одна из наиболее современных [Wolfe, 2021]. Важной составляющей данной модели является карта приоритетов, которая отражает степень активации мест в пространстве. Рассмотренные нами факторы: заметность, карта значений, контекст как глобальная структура сцены и контекст как совместная встречаемость объектов в сцене — вносят свой вклад в распределение этой активации. Внимание перемещается по карте приоритетов в порядке ее убывания. Таким образом, если ранние модели учитывали только зрительную заметность, то более современные говорят о разных источниках информации о сцене, способной направлять поиск. Однако для понимания его протекания в сценах недостаточно просто выделить разные факторы, важно также описать и механизм, стоящий за этим. Модель Вольфа позволяет это сделать.

Одним из элементов модели является долговременная память. В ней хранятся как зрительные признаки объектов, так и, судя по исследованиям структуры сцены, информация о паттернах зрительных признаков сцен и объектов, часто встречающихся друг с другом. Наибольшую активацию на карте приоритетов в таком случае получат те места, которые соотносятся по своим зрительным паттернам с целевым объектом. Возможно, что именно так это происходит при поиске в сценах, где присутствуют якорные объекты.

Появление новых факторов, объясняющих, как направляется внимание в сцене, сопровождается и появлением новых вопросов. Так, например, на сегодняшний день лучше всего изучены сцены нашего бытового обихода — кухня, ванная, офис и т. п. Однако мало что известно о структуре природных сцен. Актуален вопрос о том, можно ли перенести описанную выше иерархию объектов на них, каковы будут особенности зрительного поиск в таких сценах [Turini, Võ, 2022; Võ, 2021] и можно ли применить ее к динамичным сценам [Võ, 2021]. Дополнительно ставится вопрос о том, могут ли в зависимости от масштаба рассмотрения сцены менее крупные объекты выполнять роль якорей, которые при бóльшем масштабе выступали бы локальными [Helbing, Draschkow, Võ, 2022].

Заключение

Проведенный анализ показывает, что направление внимания в сценах зависит от нескольких факторов, которые предлагают разные объяснения для этого. С одной стороны, свой вклад вносит зрительная заметность объектов за счет определения локальных контрастов сцены. С другой — смысл и структура сцены. Первый выступает как семантическая согласованность объекта и сцены и имеет разное распределение в сцене. Вторая определяется как пространственные отношения объекта с его окружением, в частности — с другими объектами. Объединяясь вместе, данные факторы формируют карту приоритетов, задающую вниманию путь сканирования сцены.

Финансирование

Статья подготовлена в рамках выполнения научно-исследовательской работы государственного задания РАНХиГС.

Литература

Измайлов Ч.А., Черноризов А.М. Язык восприятия и мозг. Психология. Журнал Высшей школы экономики, 2005, Т. 2, No. 4, 22–52.

Фаликман М.В. Общая психология. Внимание. — М., 2006.

Biederman I., Mezzanotte R.J., Rabinowitz J.C. Scene perception: Detecting and judging objects undergoing relational violations. Cognitive Psychology, 1982, v. 14, No. 2, 143–177. DOI: https://doi.org/10.1145/1024273.1024283

Boettcher S.E.P., Draschkow D., Dienhart E., Võ M.L.-H. Anchoring visual search in scenes: Assessing the role of anchor objects on eye movements during visual search. Journal of Vision, 2018, v. 18, No. 13, 11. DOI: doi:https://doi.org/10.1167/18.13.11

Castelhano M.S., Heaven C. The relative contribution of scene context and target features to visual search in scenes. Attention, Perception and Psychophysics, 2010, v. 72, No. 5, 1283–1297. DOI: https://doi.org/10.3758/APP.72.5.1283

Castelhano M.S., Heaven C. Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. Psychonomic Bulletin and Review, 2011, v. 18, No. 5, 890–896. DOI: https://doi.org/10.3758/s13423-011-0107-8

Chan L.K.H., Hayward W.G. Visual search. WIRES Cognitive Science, 2013, v. 4, No. 4, 415–429. DOI: https://doi.org/10.1002/wcs.1235

Davenport J.L., Potter M.C. Scene Consistency in Object and Background Perception. Psychological Science, 2004,. v. 15, No. 8, 559–564. DOI: https://doi.org/10.1111/j.0956-7976.2004.00719.x

Draschkow D., Võ M.L.-H. Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. Scientific Reports, 2017, v. 7, No. 1, 16471. DOI: https://doi.org/10.1038/s41598-017-16739-x

Greene M.R. Statistics of high-level scene context. Frontiers in Psychology, 2013, v. 4. DOI: https://doi.org/10.3389/fpsyg.2013.00777

Helbing J., Draschkow D., Võ M.L.-H. Auxiliary Scene-Context Information Provided by Anchor Objects Guides Attention and Locomotion in Natural Search Behavior. Psychological Science, 2022,. v. 33, No. 9, 1463–1476. DOI: https://doi.org/10.1177/09567976221091838

Henderson J.M., Hayes T.R. Meaning-based guidance of attention in scenes as revealed by meaning maps. Nature Human Behaviour, 2017, v. 1, No. 10, 743–747. DOI: https://doi.org/10.1038/s41562-017-0208-0

Itti L., Koch C. Computational modelling of visual attention. Nat Rev Neurosci, 2001, v. 2, No. 3, 194–203. DOI: https://doi.org/10.1016/S0042-6989(99)00163-7

Mack S.C., Eckstein M.P. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. Journal of Vision, 2011, v. 11, No. 9, 9–9. DOI: https://doi.org/10.1167/11.9.9

Neider M.B., Zelinsky G.J. Scene context guides eye movements during visual search. Vision Research, 2006, v. 46, No. 5, 614–621.

Turini J., Võ M.L.-H. Hierarchical organization of objects in scenes is reflected in mental representations of objects. Scientific Reports, 2022, v. 12, No. 1, 20068. DOI: https://doi.org/10.1038/s41598-022-24505-x

Võ M.L.-H. The meaning and structure of scenes. Vision Research, 2021, v. 181, 10–20. DOI: https://doi.org/10.1016/j.visres.2020.11.003

Võ M.L.-H., Boettcher S.E., Draschkow D. Reading scenes: how scene grammar guides attention and aids perception in real-world environments. Current Opinion in Psychology, 2019, v. 29, 205–210. DOI: https://doi.org/10.1016/j.copsyc.2019.03.009

Võ M.L.-H., Wolfe J.M. Differential Electrophysiological Signatures of Semantic and Syntactic Scene Processing. Psychological Science, 2013, v. 24, No. 9, 1816–1823. DOI: https://doi.org/10.1177/095679761347695

Wolfe J.M. Guided Search 2.0 A revised model of visual search. Psychonomic Bulletin and Review, 1994, v. 1, No. 2, 202–238. DOI: https://doi.org/10.3758/BF03200774

Wolfe J.M., Alvarez G. A., Rosenholtz R., Kuzmova Y.I., Sherman A.M. Visual search for arbitrary objects in real scenes. Attention, Perception, and Psychophysics, 2011, v. 73, No. 6, 1650–1671. DOI: https://doi.org/10.3758/s13414-011-0153-3

Wolfe J.M. Guided Search 6.0: An updated model of visual search. Psychonomic Bulletin and Review, 2021, v. 28, No. 4, 1060–1092. DOI: https://doi.org/10.3758/s13423-020-01859-9

Поступила в редакцию: 25 мая 2024 г. Дата публикации: 30 декабря 2024 г.

Сведения об авторах

Кукарцев Данил Сергеевич. Магистр, факультет психологии, Институт общественных наук, Российская академия народного хозяйства и государственной службы при Президенте РФ, пр. Вернадского, 82, корп. 2, 119571, Москва, Россия.
E-mail: Kukartsev.Danil@yandex.ru

Морозов Максим Игоревич. Научный сотрудник научно-исследовательской лаборатории когнитивных исследований, факультет психологии, Институт общественных наук, Российская академия народного хозяйства и государственной службы при Президенте РФ, пр. Вернадского, 82, корп. 2, 119571, Москва, Россия.
E-mail: morozovmaksimmm@gmail.com

Ссылка для цитирования

Кукарцев Д.С., Морозов М.И. Структура реалистичных сцен как фактор, направляющий зрительный поиск. Психологические исследования. 2024. Т. 17, № 97. С. 4. URL: https://psystudy.ru

Адрес статьи:

https://doi.org/10.54359/ps.v17i97.1656

The structure of realistic scenes as a determinant guiding the visual search

Kukartsev D.S.[1], **Morozov M.I.**[1]

[1] The Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia

Visual search in scenes is guided by a variety of factors, each based on different types of scene-related information. One of these is visual salience, which relies on visual features within the scene and directs attention from the most salient areas to the least. Another factor is the scene's meaning and structure, which involves the congruence of objects with the scene they are located in. Congruence can refer to an object's high likelihood of belonging to the scene (meaning) and its spatially expected placement (structure). This paper is aimed to review these factors, identifying the limitations of early models of visual search in scenes, comparing J. Wolfe's model with recent findings on structural-meaning factors, and addressing a series of emerging theoretical questions.

Keywords: visual search, scene meaning, scene structure, scene hierarchy, anchor objects, saliency

Funding

The article was written on the basis of the RANEPA state assignment research program.

References

Biederman I., Mezzanotte R.J., Rabinowitz J.C. Scene perception: Detecting and judging objects undergoing relational violations. Cognitive Psychology, 1982, v. 14, No. 2, 143–177. DOI: https://doi.org/10.1145/1024273.1024283

Boettcher S.E.P., Draschkow D., Dienhart E., Võ M.L.-H. Anchoring visual search in scenes: Assessing the role of anchor objects on eye movements during visual search. Journal of Vision, 2018, v. 18, No. 13, 11. DOI: https://doi.org/10.1167/18.13.11

Castelhano M.S., Heaven C. The relative contribution of scene context and target features to visual search in scenes. Attention, Perception and Psychophysics, 2010, v. 72, No. 5, 1283–1297. DOI: https://doi.org/10.3758/APP.72.5.1283

Castelhano M.S., Heaven C. Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. Psychonomic Bulletin and Review, 2011, v. 18, No. 5, 890–896. DOI: https://doi.org/10.3758/s13423-011-0107-8

Chan L.K.H., Hayward W.G. Visual search. WIRES Cognitive Science, 2013, v. 4, No. 4, 415–429. DOI: https://doi.org/10.1002/wcs.1235

Davenport J.L., Potter M.C. Scene Consistency in Object and Background Perception. Psychological Science, 2004, v. 15, No. 8, 559–564. DOI: https://doi.org/10.1111/j.0956-7976.2004.00719.x

Draschkow D., Võ M.L.-H. Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. Scientific Reports, 2017, v. 7, No. 1,. 16471. DOI: https://doi.org/10.1038/s41598-017-16739-x

Falikman M.V. General psychology. Attention. — M., 2006. (in Russian)

Greene M.R. Statistics of high-level scene context. Frontiers in Psychology, 2013, v. 4. DOI: https://doi.org/10.3389/fpsyg.2013.00777

Helbing J., Draschkow D., Võ M.L.-H. Auxiliary Scene-Context Information Provided by Anchor Objects Guides Attention and Locomotion in Natural Search Behavior. Psychological Science, 2022, v. 33, No. 9, 1463–1476. DOI: https://doi.org/10.1177/09567976221091838

Itti L., Koch C.A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research, 2000, v. 40, No. 10–12, 1489–1506. DOI: https://doi.org/10.1016/S0042-6989(99)00163-7

Izmailov Ch.A., Chernorizov A.M. Language of perception and the brain. Psychology. Journal of the Higher School of Economics, 2005, v. 2, No. 4, 22-52. (in Russian)

Mack S.C., Eckstein M.P. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. Journal of Vision, 2011, v. 11, No. 9, 9–9. DOI: https://doi.org/10.1167/11.9.9

Neider M.B., Zelinsky G.J. Scene context guides eye movements during visual search. Vision Research, 2006, v. 46, No. 5, 614–621.

Turini J., Võ M.L.-H. Hierarchical organization of objects in scenes is reflected in mental representations of objects. Scientific Reports, 2022, v. 12, No. 1, 20068. DOI: https://doi.org/10.1038/s41598-022-24505-x

Võ M.L.-H. The meaning and structure of scenes. Vision Research, 2021, v. 181, 10–20. DOI: https://doi.org/10.1016/j.visres.2020.11.003

Võ M.L.-H., Boettcher S.E., Draschkow D. Reading scenes: how scene grammar guides attention and aids perception in real-world environments. Current Opinion in Psychology, 2019, v. 29, 205–210. DOI: https://doi.org/10.1016/j.copsyc.2019.03.009

Võ M.L.-H., Wolfe J.M. Differential Electrophysiological Signatures of Semantic and Syntactic Scene Processing. Psychological Science, 2013, v. 24, No. 9, 1816–1823. DOI: https://doi.org/10.1177/095679761347695

Wolfe J.M. Guided Search 2.0 A revised model of visual search. Psychonomic Bulletin and Review, 1994, v. 1, No. 2, 202–238. DOI: https://doi.org/10.3758/BF03200774

Wolfe J.M., Alvarez G.A., Rosenholtz R., Kuzmova Y.I., Sherman A.M. Visual search for arbitrary objects in real scenes. Attention, Perception, and Psychophysics, 2011, v. 73, No. 6, 1650–1671. DOI: https://doi.org/10.3758/s13414-011-0153-3

Wolfe J.M. Guided Search 6.0: An updated model of visual search. Psychonomic Bulletin and Review, 2021, v. 28, No. 4, 1060–1092. DOI: https://doi.org/10.3758/s13423-020-01859-9

Information about authors

Kukartsev Danil Sergeevich. Master, Faculty of Psychology, Institute of Social Sciences, Russian Presidential Academy of National Economy and Public Administration, Vernadsky Ave., 82, bldg. 2, 119571, Moscow, Russia.
E-mail: Kukartsev.Danil@yandex.ru

Morozov Maxim Igorevich. Research Associate of Research Laboratory of Cognitive Studies, Faculty of Psychology, Institute of Social Sciences, Russian Presidential Academy of National Economy and Public Administration, Vernadsky Ave., 82, bldg. 2, 119571, Moscow, Russia.
E-mail: morozovmaksimmm@gmail.com

For citation:

Kukartsev D.S., Morozov M.I. The structure of realistic scenes as a determinant guiding the visual search. Psikhologicheskie Issledovaniya, 2024, Vol. 17, No. 97, p. 4. https://psystudy.ru