

# КЛАСТЕРИЗАЦИЯ ЗОН КОГНИТИВНОЙ НАГРУЗКИ ПОЛЬЗОВАТЕЛЯ ПО ОКУЛОМОТОРНЫМ ДАННЫМ СЕССИЙ ПРОСМОТРА ГРАФИЧЕСКИХ СТРУКТУР

## CLUSTERING OF USER COGNITIVE LOAD ZONES BASED ON OCULOMOTOR DATA FROM GRAPHICAL STRUCTURE VIEWING SESSIONS

**B. Goryachkin  
A. Savelyev  
M. Fonin  
D. Lobanov**

*Summary.* A non-invasive approach to identifying areas of elevated cognitive load based on eye-tracking data was investigated in this paper. Gaze fixation patterns during interaction with graphical structures were analyzed. The DBSCAN density clustering method was applied to automatically detect areas of interest (AOI) without prior labeling. The Cognitive Load Index (CLI), quantifying the intensity of attention within the identified clusters, was calculated. Experiments on visual scenes of varying complexity were conducted, involving the variation of the neighborhood radius and minimum points parameters. The sensitivity of the algorithm to clusterization parameters was evaluated. The effectiveness of the method for the spatial interpretation of mental effort distribution without the use of specialized equipment was confirmed.

*Keywords:* cognitive load, eye tracking, fixations, attention zones, clustering, visualization.

**Горячкин Борис Сергеевич**

Кандидат технических наук, доцент,  
Московский государственный технический  
университет им. Н.Э. Баумана  
bsgor@mail.ru

**Савельев Алексей Александрович**

Аспирант, Московский государственный технический  
университет им. Н.Э. Баумана  
savelevvaa@mail.ru

**Фонин Максим Алексеевич**

Московский государственный технический  
университет им. Н.Э. Баумана  
max.fonin@mail.ru

**Лобанов Дмитрий Сергеевич**

Московский государственный технический  
университет им. Н.Э. Баумана  
citrusbound@mail.ru

*Аннотация.* Рассмотрен неинвазивный подход к определению областей повышенной когнитивной нагрузки на основе данных айтрекинга. Проанализированы паттерны фиксации взгляда при взаимодействии пользователя с графическими структурами. Применен метод плотностной кластеризации DBSCAN для автоматического выделения зон интереса (AOI) без предварительной разметки. Рассчитан индекс когнитивной нагрузки (CLI), количественно характеризующий интенсивность внимания в выделенных кластерах. Проведены эксперименты на визуальных сценах различной сложности с варьированием параметров радиуса окрестности и минимального числа точек. Оценена чувствительность алгоритма к настройкам алгоритма кластеризации. Подтверждена эффективность метода для пространственной интерпретации распределения умственных усилий без использования специализированного оборудования.

*Ключевые слова:* когнитивная нагрузка, отслеживание взгляда, фиксации, зоны внимания, кластеризация, визуализация.

### Введение

**К**огнитивная нагрузка отражает количество интеллектуальных ресурсов, которые человек задействует при выполнении задачи или восприятии информации. В сфере взаимодействия человека с цифровыми системами оценка когнитивной нагрузки имеет ключевое значение, поскольку позволяет выявлять элементы интерфейса или содержания, вызывающие наибольшие трудности для пользователя. Полученные сведения используются для совершенствования интерфейсов, учебных материалов, в рекламных целях и других видах визуальной информации. Важным преимуществом является

то, что оценка когнитивной нагрузки может проводиться без применения инвазивных методов, одновременно предоставляя информацию о степени концентрации и степени утомляемости пользователя.

Отслеживание траекторий взгляда и его фиксации предоставляет собой временные и пространственные данные о распределении внимания [1]. Если пользователь надолго фиксирует взгляд на конкретной области изображения, вероятнее всего это значит, что пользователь испытывает когнитивную нагрузку — либо из-за сложности воспринимаемого элемента, либо из-за его значимости. Анализируя особенности движения глаз,

можно определить участки наибольшей умственной нагрузки при выполнении задачи.

Данная работа посвящена разработке автоматизированного подхода к обнаружению зон повышенной когнитивной нагрузки на основе данных eye-tracking-системы. В статье представлены теоретические основы подхода, описана его реализация, приведены результаты экспериментальных исследований.

### Постановка задачи

Во многих практических задачах важно не только определить общий уровень когнитивной нагрузки пользователя, но и выявить локальные элементы интерфейса или фрагменты графического материала, которые вызывают наибольшие трудности. Существующие методы оценки, как правило, предоставляют усреднённый показатель за весь период взаимодействия и не позволяют выявить области, на которые приходилась какая-то часть когнитивной нагрузки и внимания [2].

В рамках данного исследования ставится задача автоматического обнаружения зон повышенной когнитивной нагрузки на основе данных движения глаз. Задача сводится к поиску кластеров точек фиксации взгляда (Aria of Interest, AOI), соответствующих областям, где пользователь задерживается дольше или чаще, что указывает на повышение умственных усилий. Реализация задачи включает несколько этапов: сбор данных о движении глаз, выделение фиксаций, их последующую кластеризацию и оценку степени нагрузки для каждой обнаруженной зоны. Необходимо спроектировать и разработать алгоритм определения степени когнитивной нагрузки для зарегистрированных окуломоторных данных, а также способ кластеризации и визуализации анализируемых данных для их наглядного, графического представления.

Реализация поставленной задачи потенциально позволит выявлять объекты и зоны интереса пользователей, или же элементы интерфейса, вызывающие повышенную когнитивную нагрузку для их оценки [3] и улучшения.

### Когнитивная нагрузка как базовый индикатор интеллектуальной деятельности пользователя компьютера

В рамках теории когнитивной нагрузки выделяют несколько категорий умственных усилий, которые совместно формируют общий уровень интеллектуальной нагрузки пользователя [4]. Традиционно различают три основные разновидности:

1. Внутренняя когнитивная нагрузка определяется сложностью содержания: объемом информа-

ции, числом взаимосвязей между элементами, абстрактностью и трудностью рассматриваемых концепций. Этот тип нагрузки является неизбежным, поскольку обусловлен самим содержанием материала и не зависит от того, как он представлен.

2. Внешняя когнитивная нагрузка возникает из-за особенностей подачи информации — перегруженного интерфейса, лишних деталей, отвлекающих факторов. Она считается нежелательной, поскольку является следствием внешних факторов и затрудняет восприятие.
3. Смысловая (или продуктивная) когнитивная нагрузка отражает усилия, направленные на понимание, осмысление и усвоение новой информации. Этот тип нагрузки считается полезным, поскольку связан с активным формированием знаний и пониманием материала.

В совокупности эти три компонента определяют общий уровень когнитивной нагрузки, результаты анализа которого можно использовать для оптимизации и улучшения интерфейсов или графических структур, которые в свою очередь используются для обучения или работы, что в конечном итоге приведет к повышению эффективности восприятия.

Также выделяют 3 наиболее распространённые группы методов оценки нагрузки: физиологические, поведенческие и субъективные [5].

1. Физиологические методы основываются на анализе биосигналов, чувствительных к уровню умственного напряжения: частоты сердечного ритма, кожно-гальванической реакции, данных ЭЭГ, диаметра зрачков, частоты моргания и других параметров. Эти методы обеспечивают объективные данные, но требуют специализированного оборудования и часто оказываются неудобными в применении.
2. Поведенческие методы оценивают нагрузку по особенностям выполнения задачи и по параметрам движений глаз. Увеличение длительности выполнения, рост количества ошибок, особенности фиксаций и саккад, а также повторные возвраты взгляда могут указывать на рост когнитивной нагрузки. Поскольку этот подход не требует вмешательства в деятельность пользователя и может быть реализован с помощью трекаera взгляда, он является основой разработанного в данной работе подхода.
3. Субъективные методы предполагают самооценку пользователем своего состояния через опросы и шкалы, такие как NASA TLX. Они просты в применении, но субъективны и не позволяют отслеживать динамику нагрузки в реальном времени, поскольку обычно применяются после выполнения задания.

В данном исследовании основное внимание уделяется поведенческому подходу, в частности анализу глазодвигательных данных, который является неинвазивным, обладает высокой временной и пространственной точностью.

**Обработка данных трекера взгляда и реализация метода автоматического выделения зон внимания с уровнем умственной нагрузки**

В исследовании применяется самописный трекер взгляда [6] и его различные модули для визуализации данных поведения взгляда. Схема работы системы представлена на рисунке 1 и функционирует следующим образом: изображение пользователя захватывается веб-камерой компьютера, после чего происходит анализ лица, определяется положение зрачков. На основе изменений положения зрачков относительно калибровочных точек вычисляется зона взгляда на экране.

Система отслеживания взгляда предоставляет исходные окуломоторные данные для анализа, а именно координаты точек взгляда на экране и его фиксации. На основе этих данных формируются различные визуализации, представленные на рисунке 2, среди которых диаграммы разброса регистрируемых покадрово точек

взгляда и тепловые карты, отражающие насыщенностью цвета интенсивность и продолжительность фиксации.

В нашей работе мы автоматически выделяем зоны визуального интереса пользователя (AOI), не прибегая к их предварительной ручной разметке, которая не всегда может быть объективна. Мы используем собранные глазодвигательные данные для выявления истинных зон интереса пользователя на конкретном изображении в конкретной сессии отслеживания взгляда. Для этого применяется алгоритм DBSCAN — метод выделения плотностных структур, способный формировать кластеры произвольной формы без необходимости заранее задавать их количество [7]. Его параметрами являются радиус окрестности  $\epsilon$ , определяющий зону поиска соседних точек, и минимальное число точек MinPts, необходимое для формирования кластера. Варьирование этих параметров позволяет определять различные зоны интереса на разном уровне масштабирования: глобальные и локальные. Диапазон для эффективного поиска кластеров у окрестности  $\epsilon$  от 10 до 100 пикселей, у MinPts от 5 до 30 штук. В таблице 1 приведены примеры вариации параметров алгоритма с описаниями ожидаемых результатов его работы с окуломоторными данными, используемыми в данной работе.

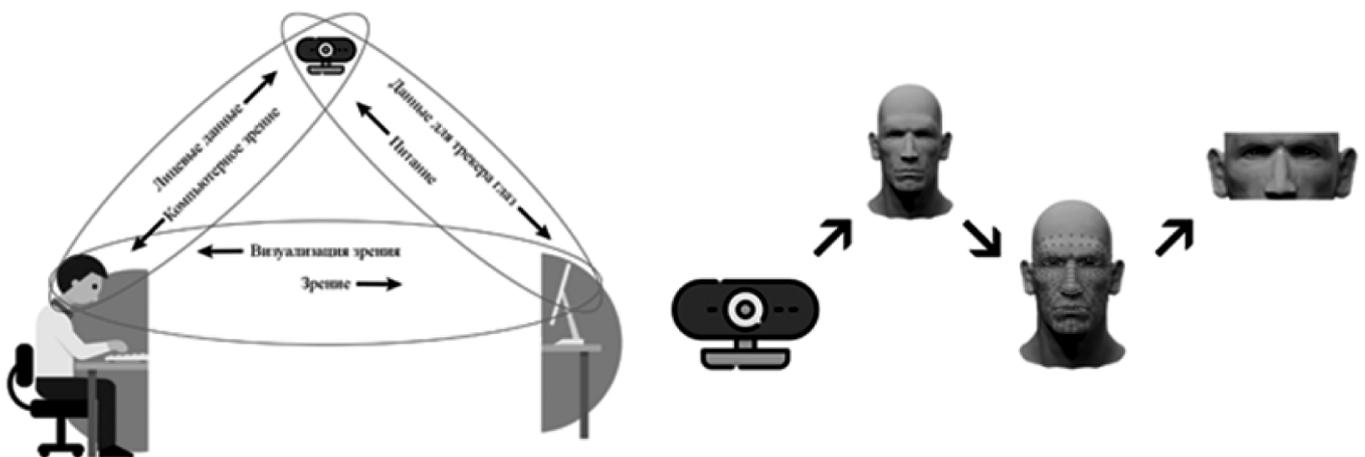


Рис. 1. Общий алгоритм работы системы

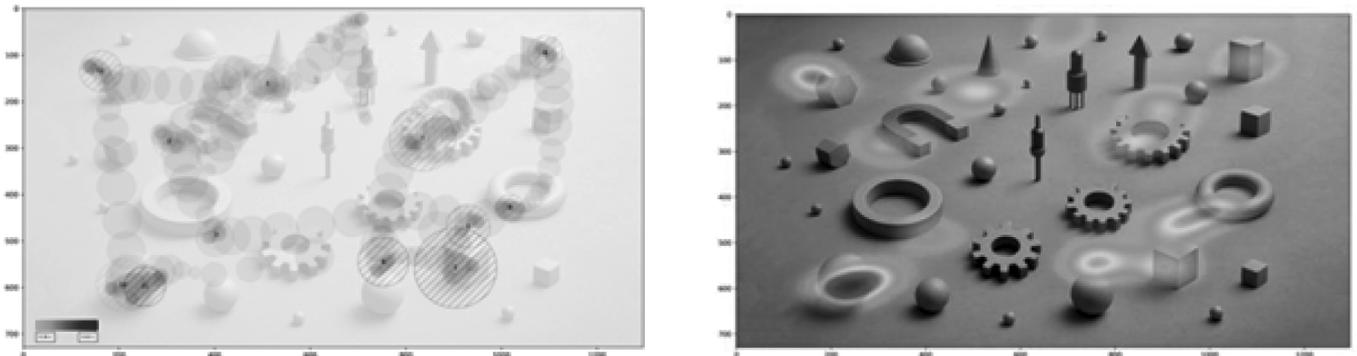


Рис. 2. Диаграмма рассеивания и тепловая карта точек взгляда

Таблица 1.  
Варьирование значений параметров DBSCAN

Значения параметров	Результат
$< \epsilon, < \text{MinPts}$	Алгоритм ищет небольшие скопления. Получаем много мелких, дробных кластеров
$< \epsilon, > \text{MinPts}$	Алгоритм ищет точки с высокой с высокой пространственной плотностью.
$> \epsilon, < \text{MinPts}$	Алгоритм строит кластеры по далеким точкам с низкой пространственной плотностью
$> \epsilon, > \text{MinPts}$	Алгоритм объединяет широкие области, но требует, чтобы в них было большое количество точек. Получаем крупные зоны общего интереса

После выделения кластеров необходимо определить, как именно взаимодействовал пользователь с визуальными объектами под ними. Инструментарий, используемый для сбора данных, неинвазивный, поэтому мы не можем использовать нейрофизиологические методы для определения уровня умственных усилий. Было принято решение использовать композитный индекс когнитивной нагрузки (Cognitive Load Index, CLI), представляющий собой взвешенную линейную комбинацию двух параметров, извлекаемых из данных фиксации взгляда в пределах кластера (уравнение 1).

$$CLI = \omega_1 \cdot DwellK_i + \omega_2 \cdot AvgFK_i \quad (1)$$

В основе рассматриваемого метода лежит гипотеза «Глаз-Разум» (Eye-Mind Hypothesis [8]), сформулированная Джастом и Карпентером в 1980 году, согласно которой существует тесная временная и пространственная корреляция между поведением взгляда и обрабатываемой мозгом информацией. Предложенная формула 1 решает задачу определения умственной нагрузки через интеграцию двух параметров: суммарного времени фиксации и средней продолжительности фиксации для соответствующих кластеров.

Параметр ***DwellK<sub>i</sub>***, (**мера распределения внимания**) — нормированное суммарное время фиксации в *i*-ом кластере, представляющее собой долю общего времени просмотра, затраченного на анализ данной области. Вычисляется как отношение суммарной длительности всех фиксации, попавших в кластер, к общему времени всех фиксации в сессии. Данный параметр отражает относительную значимость области для пользователя и служит базовым индикатором распределения внимания.

Параметр ***AvgFK<sub>i</sub>***, (**мера глубины обработки**) — среднее время одиночной фиксации в кластере, характеризующее глубину (интенсивность) когнитивной

обработки информации. Вычисляется как среднее арифметическое длительностей всех фиксации, принадлежащих кластеру. Это позволяет различать ситуации с одинаковым суммарным временем, но разной природой взаимодействия. Зона, просмотренная бегло с множеством коротких фиксации, будет иметь меньший CLI, чем зона, изученная пристально, с несколькими, но длинными фиксациями, даже если суммарное время одинаково. Это критически важно для выявления именно нагрузки, а не просто интереса.

Вопрос определения весовых коэффициентов является центральным для валидности модели. Были выбраны значения  $\omega_1 = 0.6$ ,  $\omega_2 = 0.4$ , с приоритетом у первого параметра, и вот почему:

1. Суммарное время фиксации является характеристикой, которая более устойчива к шуму, чем среднее значение. Если дать ему меньший вес, индекс CLI станет слишком чувствительным к коротким эпизодам «задумчивости» в неважных зонах.
2. Средняя длительность фиксации является метрикой с высокой дисперсией. Так одна случайная длинная фиксация, например, когда пользователь отвлекся, глядя в точку, может резко завесить среднее.
3. В многокритериальном анализе часто используется разбиение 60/40 для баланса между «Количеством» (общее время) и «Качеством» (интенсивность фиксации) [9].

В итоге мы получаем карту когнитивной нагрузки, на которой кластеры, у которых значение CLI больше, чем у других, идентифицируются как зоны, потребовавшие больших интеллектуальных усилий при визуальном взаимодействии, что может быть обусловлено как высокой информационной значимостью элементов, так и сложностью их обработки. Таким образом описанный подход обеспечивает автоматизированное измерение когнитивной нагрузки на основе поведенческих данных, доступных при использовании неинвазивной системы айтрекинга.

### Экспериментальная часть

Для тестирования расчета CLI и влияния параметров кластеризации на выделение зон интереса, был проведен ряд экспериментов, выполненных на различных визуальных сценах. В каждом случае параметры алгоритма DBSCAN изменялись, что позволило получить различающиеся по структуре и детализации результаты кластеризации и оценить чувствительность метода к выбираемым настройкам.

Первый эксперимент с высоко-детализированными объектами представлен на рисунке 3. В верхнем ряду приводятся зарегистрированные данные взгляда, ниж-

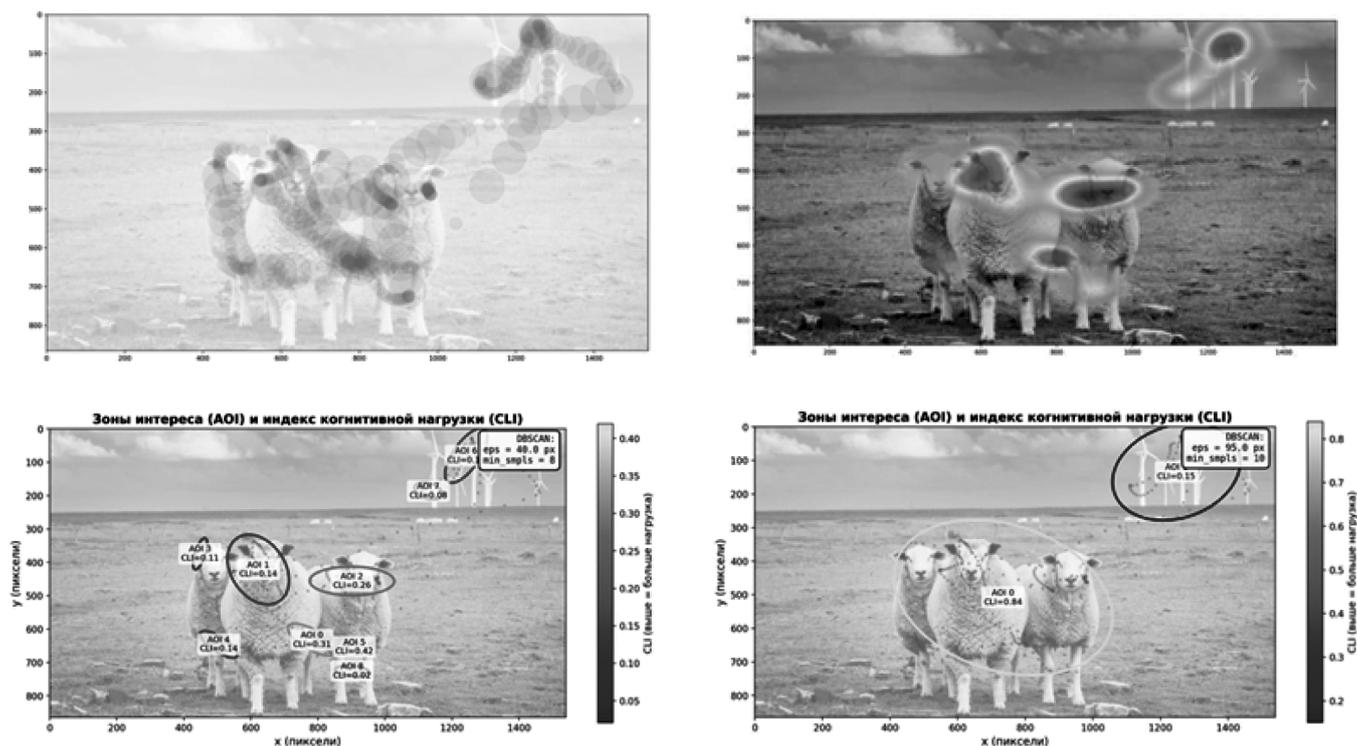


Рис. 3. Первая сцена с высоко-детализированными объектами

няя часть в свою очередь демонстрирует выявленные зоны внимания с когнитивной нагрузкой при разных параметрах алгоритма кластеризации. Полученные значения когнитивной нагрузки приведены в таблице 2.

Таблица 2.

CLI для зон интереса высоко-детализированных объектов

CLI зон интереса с параметрами DBSCAN eps=40, min_samples=8								
AOI <sub>0</sub>	AOI <sub>1</sub>	AOI <sub>2</sub>	AOI <sub>3</sub>	AOI <sub>4</sub>	AOI <sub>5</sub>	AOI <sub>6</sub>	AOI <sub>7</sub>	AOI <sub>8</sub>
0.31	0.14	0.26	0.11	0.14	<b>0.42</b>	0.10	0.08	0.02
CLI зон интереса с параметрами DBSCAN eps=95, min_samples=10								
AOI <sub>0</sub>					AOI <sub>1</sub>			
<b>0.84</b>					0.15			

Во втором эксперименте пользователь смотрел на изображение с несколькими геометрическими объектами (см. рисунок 4), на которые примерно в равной мере было распределено его внимание. Результаты произведенных расчетов при различных параметрах кластеризации представлены в таблице 3.

В третьем эксперименте, представленном на рисунке 5, пользователю демонстрировалось изображение с одним монотонным рисунком (паттерном), не содержащим явно выраженных семантических объектов. По сути, это визуальный шум, не имеющий явных смысловых акцентов. Данные окулomotorной активности демонстри-

руют, что фиксации взгляда распределены достаточно разрозненно, а выраженные кластеры интереса отсутствуют. Несмотря на попытку алгоритма DBSCAN сформировать зоны внимания, полученные кластеры не сопоставляются с конкретными смысловыми объектами под ними, за исключением таковых, а значения индекса CLI оказываются низкими и близкими друг к другу.

Наконец в рамках четвертого эксперимента (см. рисунок б), пользователю предлагалось визуально изучить сложную композицию, сформированную из набора разнообразных паттернов. Алгоритм определения зон интереса и индекса когнитивной нагрузки в данном случае позволяет оценить распределение пользовательского внимания и демонстрирует то, что в случае с набором различных паттернов пользователь проявляет повышенную умственную деятельность, выражающуюся в поведении взгляда.

### Обсуждение результатов

Анализ результатов четырех экспериментов позволяет выявить ключевые закономерности работы предложенного подхода оценки когнитивной нагрузки. Результаты подтверждают, что интеграция алгоритма DBSCAN с композитным индексом CLI формируют гибкую модель определения визуального внимания, чувствительную к пространственной составляющей исходных данных.

Первый и второй эксперименты ярко демонстрирует выраженную зависимость детализации карт когнитив-

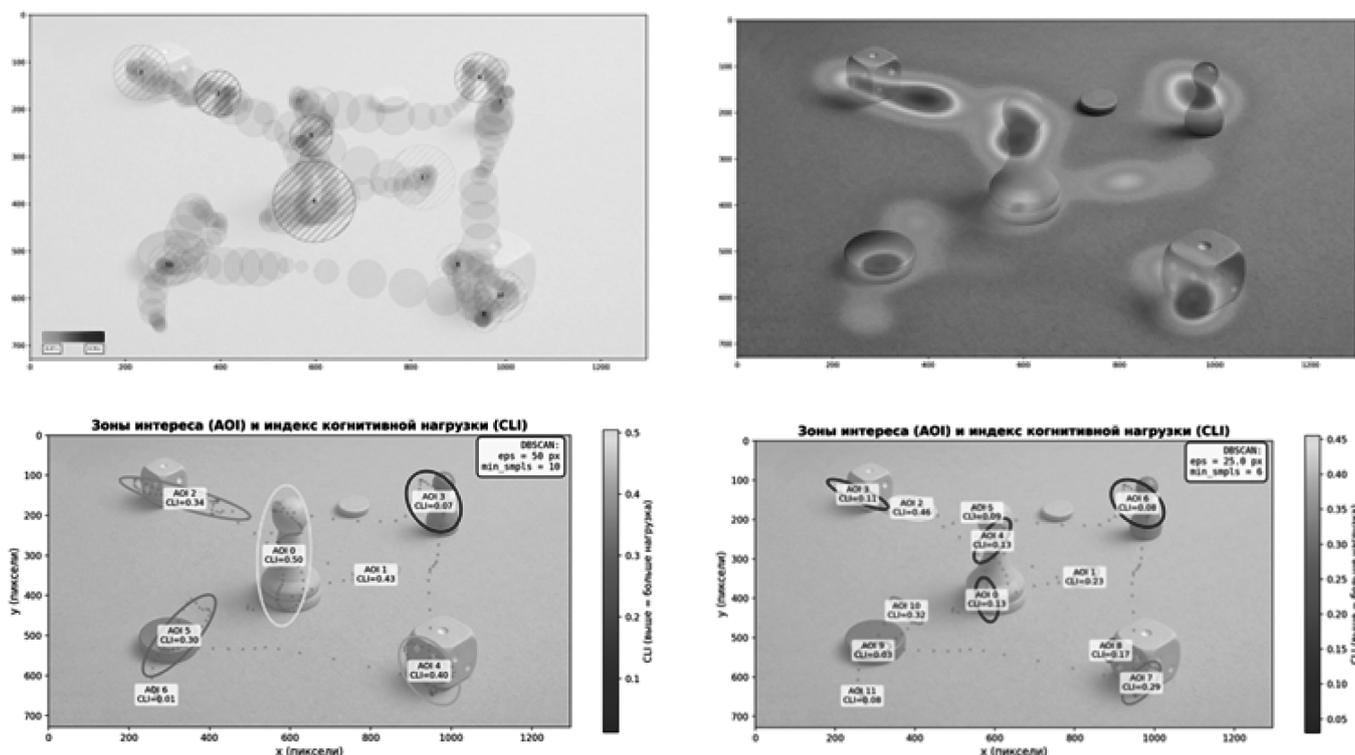


Рис. 4. Вторая сцена с различными геометрическими объектами

Таблица 3.

CLI для зон интереса геометрических фигур

CLI зон интереса с параметрами DBSCAN eps=95, min_samples=10						
AOI <sub>0</sub>	AOI <sub>1</sub>	AOI <sub>2</sub>	AOI <sub>3</sub>	AOI <sub>4</sub>	AOI <sub>5</sub>	AOI <sub>6</sub>
0.5	0.43	0.34	0.07	0.4	0.3	0.01

CLI зон интереса с параметрами DBSCAN eps=25, min_samples=10											
AOI <sub>0</sub>	AOI <sub>1</sub>	AOI <sub>2</sub>	AOI <sub>3</sub>	AOI <sub>4</sub>	AOI <sub>5</sub>	AOI <sub>6</sub>	AOI <sub>7</sub>	AOI <sub>8</sub>	AOI <sub>9</sub>	AOI <sub>10</sub>	AOI <sub>11</sub>
0.13	0.23	0.46	0.11	0.13	0.09	0.08	0.29	0.17	0.03	0.32	0.08

ной нагрузки от настроек радиуса окрестности и плотности точек кластеров. Малые значения параметров способствуют выявлению локальных зон интереса, позволяя определять небольшие события когнитивной обработки, связанные с отдельными элементами сцены. Увеличение значений параметров приводит к обобщению зон, объединяя разрозненные фиксации в большие кластеры. Свойство масштабируемости позволяет применять данное ПО как на уровне детального изучения элементов интерфейсов, так и для оценки общей композиционной структуры сцены.

В условиях отсутствия семантической структуры в третьем эксперименте распределение фиксаций носит стохастический характер, что приводит к формированию случайных кластеров с низким значением значений

CLI, что отражено в таблице 4. Напротив, наличие выраженной смысловой нагрузки в четвертом эксперименте, в котором пользователь изучал разнообразные паттерны, сопоставляя между собой, выступает катализатором когнитивной активности, значения которой приведены в таблице 5. Реализованный подход подтвердил свою валидность в задачах по исследованию взаимодействия человека с графическими материалами.

### Заключение

В ходе выполненного исследования был разработан и экспериментально протестирован подход автоматического обнаружения областей повышенной когнитивной нагрузки на основе данных движения глаз. Мы использовали самописный трекер взгляда для сбора данных окулomotorной активности пользователя во время изучения разнообразных изображений. Основой подхода стала кластеризация фиксаций с использованием алгоритма DBSCAN и вычисление композитного показателя CLI, позволяющего количественно оценивать степень вовлеченности пользователя в анализ различных частей визуальной сцены. Результаты экспериментов подтвердили эффективность предложенного решения при наличии достаточного объема данных и показали пригодность использования в подобных исследованиях. Реализованная программа обеспечивает пространственную интерпретацию распределения внимания, что делает его полезным инструментом при изучении интерфейсов и визуальных материалов для их оптимизации

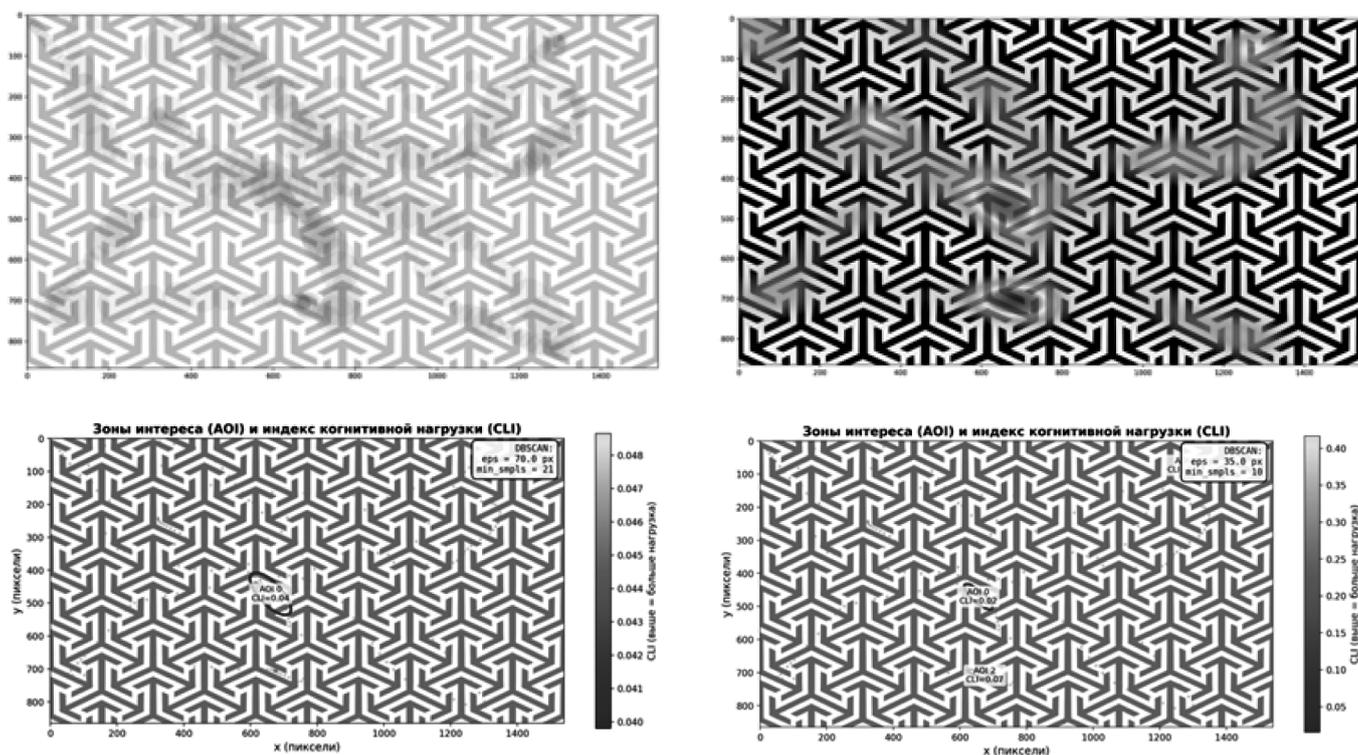


Рис. 5. Третья сцена с одним паттерном

Таблица 4.

CLI для зон интереса одного паттерна

CLI зон интереса с параметрами DBSCAN eps=70, min_samples=21	
<b>AOI<sub>0</sub></b>	
<b>0.04</b>	
CLI зон интереса с параметрами DBSCAN eps=35, min_samples=10	
AOI <sub>0</sub>	AOI <sub>1</sub>
0.02	<b>0.07</b>

ЛИТЕРАТУРА

- Lai M.L., et al., A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educational Research Review, 2013, vol. 10, pp. 90–115. DOI: 10.1016/J.EDUREV.2013.10.001.
- Paas F., Van Merriënboer, J.J.G., Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 2019, vol. 32, pp. 261–292. DOI: 10.1007/s10648-019-09465-5.
- Методика оценки удобства интерфейса веб-приложения на основе количественных критериев / С.А. Соловьев, С.С. Гончаров, А.А. Мартынова, Б.С. Горячкин // Естественные и технические науки. — 2022. — № 4(167). — С. 199–205. — DOI 10.25633/ETN.2022.04.17. — EDN XAMDYR.
- Sweller J., Ayres P., Kalyuga S., Cognitive Load Theory. Springer, 2011. DOI: 10.1007/978-1-4419-8126-4.
- Paas F.G.W.C., Van Merriënboer, J.J.G., Instructional control of cognitive load in the training of complex cognitive tasks. Educational Psychology Review, 1994, vol. 6, no. 4, pp. 351–371. DOI: 10.1007/BF02213420.
- Горячкин Б.С. Определение акцентов внимания с помощью интеллектуального самописного трекера глаз / Б.С. Горячкин, А.А. Савельев // Computational Nanotechnology. — 2024. — Т. 11, № 4. — С. 35–44. — DOI 10.33693/2313-223X-2024-11-4-35-44. — EDN FYRWAJ.
- Ester M. et al., A density-based algorithm for discovering clusters in large spatial databases with noise // kdd. — 1996. — V. 96. — №. 34. — P. 226–231.
- Just M.A., Carpenter P.A., A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 329–354. DOI:10.1037/0033-295X.87.4.329

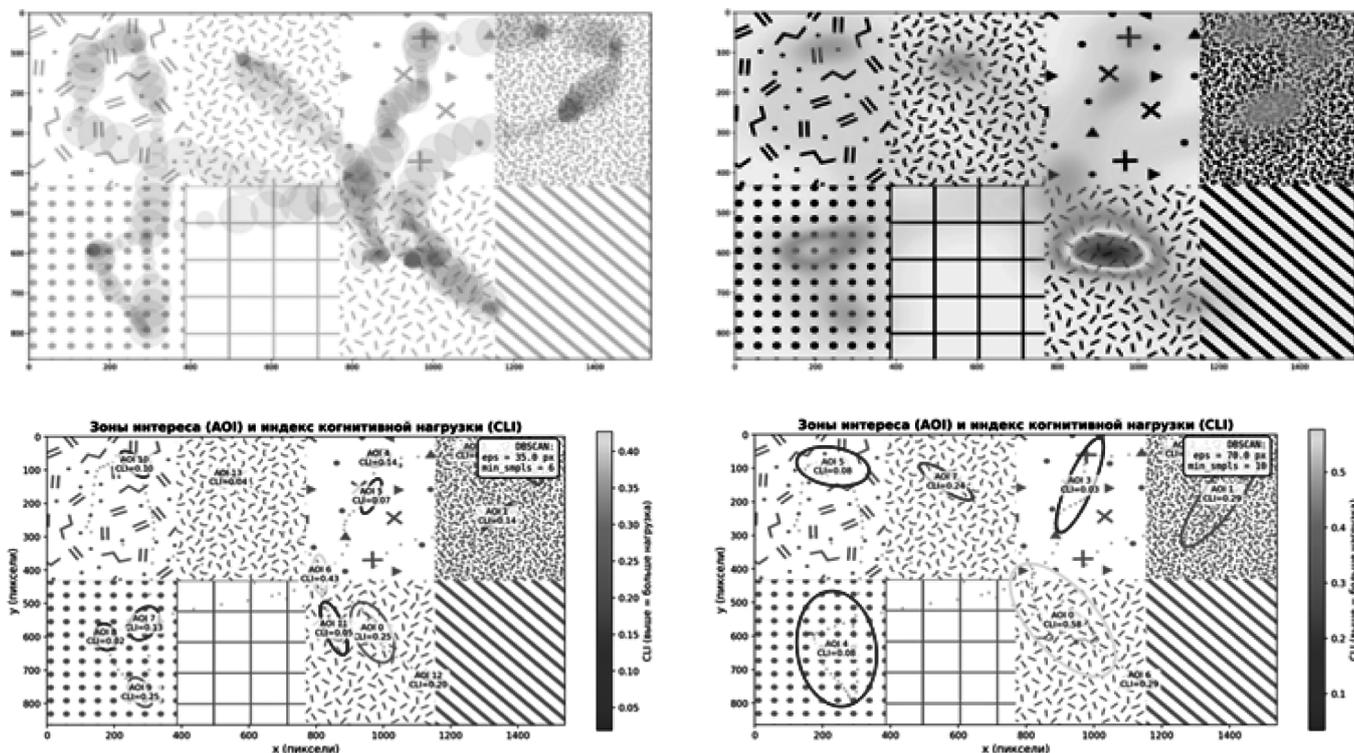


Рис. 6. Четвертая сцена с композицией паттернов

Таблица 5.

CLI для зон интереса композиции паттернов

CLI зон интереса с параметрами DBSCAN eps=35, min_samples=6													
AOI <sub>0</sub>	AOI <sub>1</sub>	AOI <sub>2</sub>	AOI <sub>3</sub>	AOI <sub>4</sub>	AOI <sub>5</sub>	AOI <sub>6</sub>	AOI <sub>7</sub>	AOI <sub>8</sub>	AOI <sub>9</sub>	AOI <sub>10</sub>	AOI <sub>11</sub>	AOI <sub>12</sub>	AOI <sub>13</sub>
0.25	0.14	0.18	0.16	0.14	0.07	<b>0.43</b>	0.13	0.02	0.25	0.1	0.05	0.2	0.04

CLI зон интереса с параметрами DBSCAN eps=70, min_samples=10							
AOI <sub>0</sub>	AOI <sub>1</sub>	AOI <sub>2</sub>	AOI <sub>3</sub>	AOI <sub>4</sub>	AOI <sub>5</sub>	AOI <sub>6</sub>	AOI <sub>7</sub>
<b>0.58</b>	0.29	0.28	0.31	0.03	0.08	0.29	0.24

- Meghanathan R.N., et al. Fixation duration surpasses pupil size as a measure of memory load. *Frontiers in Psychology*, 2015, vol. 6, article 1063. DOI: 10.3389/fpsyg.2015.01063.
- Горячкин Б.С. Эффективность принципов адаптивной верстки при разработке пользовательских интерфейсов / Б.С. Горячкин, Д.В. Бобров // *Динамика сложных систем — XXI век. — 2023. — Т. 17, № 1. — С. 55–62. — DOI 10.18127/j19997493-202301-04. — EDN SEJPOZ.*

© Горячкин Борис Сергеевич (bsgor@mail.ru); Савельев Алексей Александрович (savelevvaa@mail.ru);  
Фонин Максим Алексеевич (max.fonin@mail.ru); Лобанов Дмитрий Сергеевич (citrusbound@mail.ru)  
Журнал «Современная наука: актуальные проблемы теории и практики»