

УДК 004.932

Повышение эффективности анализа когнитивного восприятия сложных графических изображений за счет расширения функционала самописного трекера глаз

Горячкин Борис Сергеевич^(*)

bsgor@bmstu.ru

Цибизова Татьяна Юрьевна

mumc@bmstu.ru

Савельев Алексей Александрович

savelevaa@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Россия

Аннотация. В данной статье представлен подход к расширению функционала системы трекинга взгляда с целью повышения эффективности анализа когнитивного восприятия сложных графических изображений. Внедрение механизма вычисления фиксации взгляда позволяет точно определять продолжительность концентрации пользователя на определенных точках и характер его взгляда. Также улучшен UI/UX интерфейс за счет добавления указателя взгляда, отображающего направление взгляда, выходящего за пределы видимой зоны экрана. Дополнительно введена возможность подсчета морганий пользователя с использованием метрики EAR. В статье также обсуждаются преимущества использования диаграммы рассеивания для визуализации данных взгляда по сравнению с традиционной тепловой картой.

Ключевые слова: зрение, взгляд, отслеживание, фиксация, моргание, диаграмма рассеивания, тепловая карта, UI/UX.

Enhancing the efficiency of cognitive perception analysis of complex graphical images through expanding the functionality of a custom eye tracker

Goryachkin Boris Sergeevich^(*)

bsgor@bmstu.ru

Tsibizova Tatiana Yuryevna

mumc@bmstu.ru

Savelyev Alexey Alexandrovich

savelevaa@student.bmstu.ru

BMSTU, Moscow, Russia

Abstract. This article presents an approach to expanding the functionality of an eye-tracking system to enhance the efficiency of cognitive perception analysis for complex graphical images. The integration of a fixation calculation mechanism allows for precise determination of how long a user concentrates on specific points and the nature of their gaze. The UI/UX has also been improved by adding a gaze pointer, which tracks the direction of the gaze when it moves beyond the visible screen area. Additionally, a feature to count user blinks using the EAR metric has been introduced. The article also discusses the advantages of using scatter plot visualization for gaze data compared to traditional heat maps.

Keywords: vision, gaze, tracking, fixation, blink detection, scatter plot, heat map, UI/UX

Введение. Наш взгляд — это инструмент восприятия, с помощью которого мы можем взаимодействовать с окружающим миром. Паттерн поведения взгляда человека наполнен мелкими движениями глаз и их фиксациями. Наблюдение за этим процессом может многое рассказать о том, как мы думаем, воспринимаем и реагируем на сложные образы. Это позволит изучить нюансы зрительного восприятия, раскрыть закономерности, которые лежат в основе внимания и восприятия сложных визуальных объектов [1].

Традиционные инструменты отслеживания взгляда зачастую не имеют подобного функционала, который позволил бы улавливать не просто направление, но и характер взгляда. Представьте трекер, который не просто фиксирует точку внимания, но анализирует, почему и насколько долго взгляд задержался на одном объекте, позволяет зафиксировать незаметные движения и оценить ритм морганий — все те аспекты, которые ускользают от стандартных систем. Мы расширили возможности самописного трекера, добавив механизмы для регистраций характеристик взгляда для более точного последующего его анализа.

Материалы и методы. Система отслеживания взгляда реализует в себе интерполяцию координат зрачков в координаты взгляда на экране с помощью двух калибровочных координат — граничных значений зрачков [2]. Выражения расчета координат представлены уравнениями

$$X_{\text{gaze}} = \frac{(X_{\text{pulp}} - X_{\text{top_left}})}{X_{\text{bottom_right}} - X_{\text{top_left}}} \cdot \text{Width}; \quad (1)$$

$$Y_{\text{gaze}} = \frac{(Y_{\text{pulp}} - Y_{\text{top_left}})}{Y_{\text{bottom_right}} - Y_{\text{top_left}}} \cdot \text{Height}, \quad (2)$$

где $X_{\text{gaze}}, Y_{\text{gaze}}$ — рассчитываемые координаты взгляда, $X_{\text{pulp}}, Y_{\text{pulp}}$ — координаты зрачков пользователя; $X_{\text{top_left}}, Y_{\text{top_left}}$ — координаты зрачков пользователя, при направлении взгляда в левый верхний угол; $X_{\text{bottom_right}}, Y_{\text{bottom_right}}$ — координаты зрачков пользователя, при направлении взгляда в правый нижний угол; $\text{Width}, \text{Height}$ — ширина и высота экрана соответственно.

Обобщенный алгоритм работы (поэтапно) этой системы можно записать следующим перечислением:

- 1) калибровка;
- 2) получение данных со средства видеозаписи;
- 3) определение глаз человека;
- 4) сбор координат зрения;
- 5) получение соотношения координат зрения с объектом интереса;
- 6) визуализация данных, полученных в результате соотношения.

Написанный на языке программирования Python, трекер покадрово обрабатывает данные с веб-камеры пользователя используя модуль OpenCV. Координаты зрачков трекер получает от модуля mediapipe в виде сетки точек лица (FaceMesh), из которой выбирались необходимые для расчета. Описанный алгоритм изображен на рис. 1.

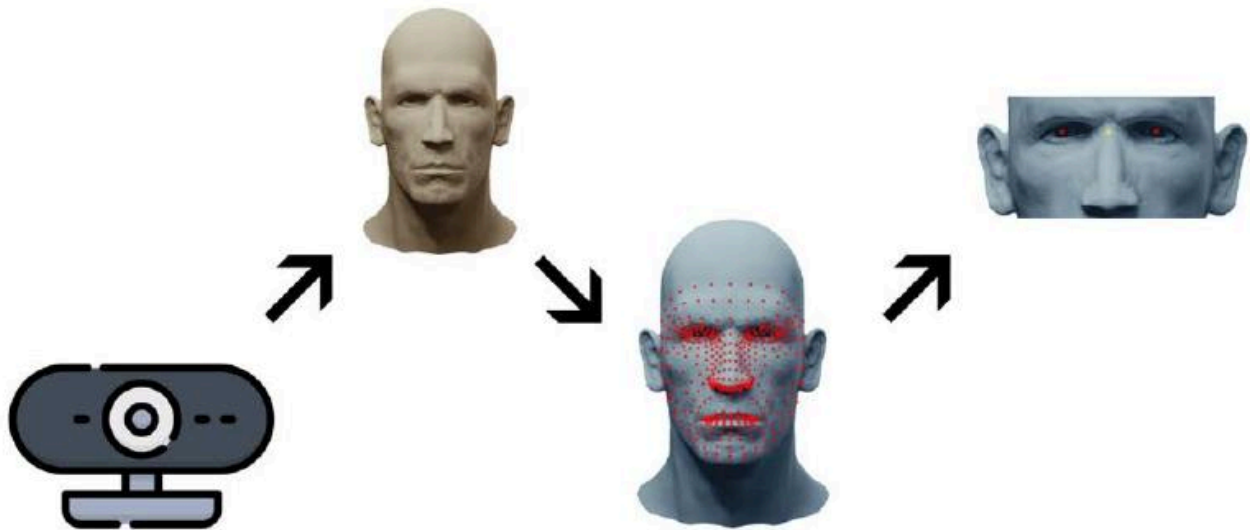


Рис. 1. Сбор и анализ данных лица пользователя

К полученным расчетным координатам применяются шумовые фильтры и определенная обработка данных, для более точной и плавной визуализации зоны взгляда пользователя на экране [3].

Также организованы подготовка и сохранение данных взгляда для дальнейшего построения по ним диаграммы рассеивания (рис. 2) с полупрозрачными окружностями, отражающими точки взгляда пользователя [4].

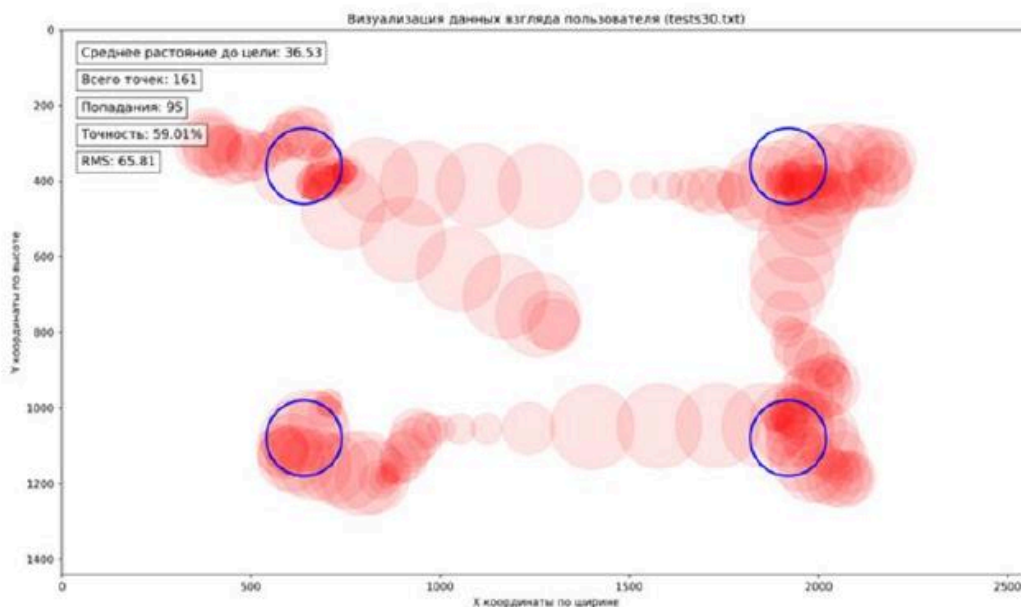


Рис. 2. Диаграмма рассеивания

Диаграмму рассеивания можно считать аналогом тепловой карты, по которой можно понять на какие из объектов на экране смотрел пользователь и определить объекты интереса среди них.

Результаты. Фиксация — это состояние, когда взгляд стабильно удерживается в одной области экрана, в отличие от быстрого перемещения взгляда, называемого саккадой. Время фиксации взгляда у человека варьируется в диапазоне от 100 до 600 мс. Для реализации данного функционала пороговое значение для регистрации фиксации было выбрано 300 мс, как середина диапазона.

Второе ключевое пороговое значение для фиксаций — это максимальный радиус области фиксации [5]. Оно установлено как 150 пикселей — это примерно 1.5 размера от фовеального зрения (1 градуса области зрения), которое мы вычислили как 101 пикселей для экрана 13 дюймов с разрешением 2560 на 1440 пикселей при удалении глаз на 65 сантиметров. Расчет представлен уравнениями

$$d = 89,35 \text{ px/см}, l = 2 \times 65 \times \tan\left(\frac{\theta}{2}\right) = 1,135 \text{ см}, \quad (3)$$

$$R = 89,35 \times 1,135 = 101,4 \text{ px},$$

где d — плотность пикселей на сантиметр экрана, l — длина участка фовеального зрения ($\theta = 1$), R — радиус области фиксации взгляда.

Выбор порога в 1.5 размера фовеального зрения объясняется необходимостью учета небольших движений головы и глаз пользователя, которые могут происходить, пока он сфокусирован на одном объекте. Фиксация не обязательно строго ограничена областью в 1° , поскольку даже при удержании взгляда на определенной точке могут происходить малые перемещения взгляда внутри области. Это своего рода «зона допуска», позволяющая считать фиксацией удержание взгляда на области.

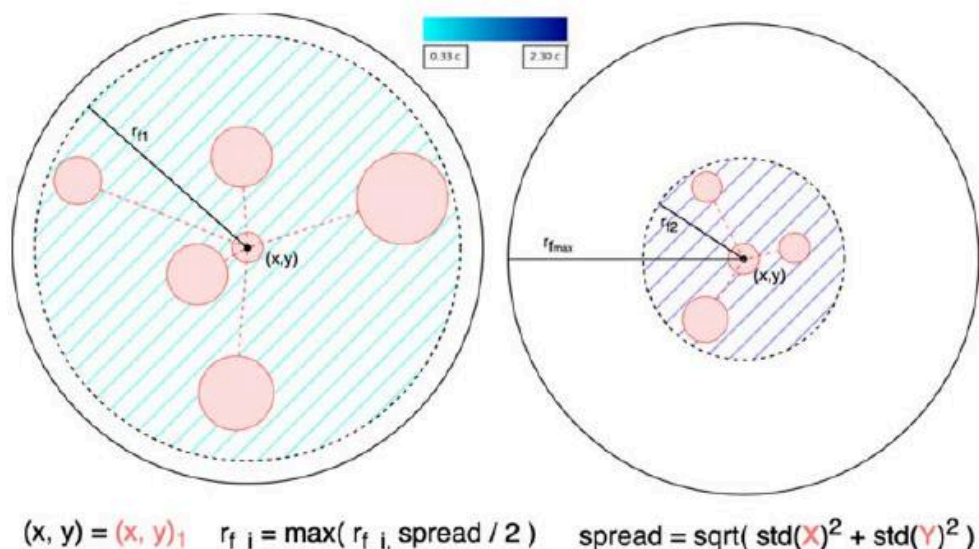


Рис. 3. Расчет фиксаций взгляда

На рис. 3 представлен пример расчета фиксаций взгляда. В случае если фиксация взгляда перешагнула за пороговое значение времени, мы начинаем

отслеживание фиксации, отслеживания как именно взгляд вел себя. Чем больше разброс точек взгляда внутри фиксации, тем больше будет ее окружность в последующем на визуализации. Чем дольше длилась фиксация по времени, тем более темнее будет ее синий цвет.

Когда пользователь использует трекер взгляда, основной задачей является отслеживание того, куда направлен взгляд, и отображение его на экране. Однако бывают ситуации, когда взгляд выходит за пределы экрана в случае смещения пользователя от места калибровки. Чтобы решить эту проблему, реализован указатель направления взгляда — небольшой круг на границе экрана, который указывает направление, в котором «потерялась» точка взгляда [6]. На рисунке 4 показаны расчеты координат на границе экрана, а также как определяется угол и расстояние для указания направления.

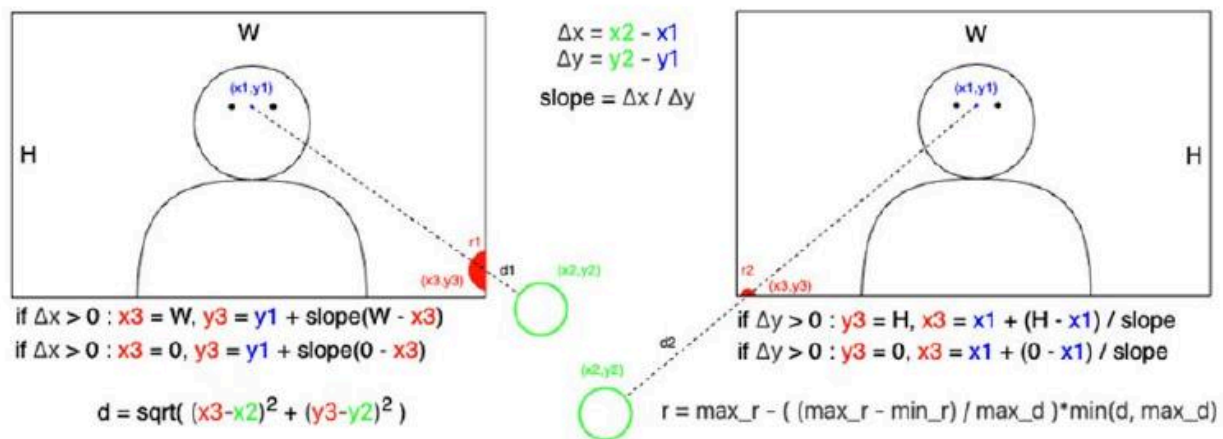


Рис. 4. Расчет параметров указателя направления взгляда

Когда взгляд уходит за границы экрана, система трекера вычисляет, на какой именно стороне экрана должно появиться указание направления. Этот расчет происходит на основе определения места пересечения границы экрана (по горизонтали или по вертикали) и угла наклона прямой от глаз пользователя до точки взгляда.

Размер указателя взгляда зависит от того, на сколько рассчитываемый взгляд пользователя удален от рабочей зоны экрана: чем дальше взгляд, тем меньше радиус красного указателя взгляда.

Адаптивный размер указателя, меняющийся в зависимости от удаленности взгляда, решает несколько важных задач:

Улучшение UI: лаконичное решение проблемы потерявшейся точки взгляда, не перегружающее графический интерфейс системы.

Улучшение UX: Изменение размера указателя в зависимости от удаленности взгляда позволяет пользователю сразу понять, где находится его взгляд, без необходимости обдумывать каждый раз, как далеко он от экрана. Это снижает когнитивную нагрузку и делает взаимодействие с трекером естественным и удобным.

Этот модуль отслеживает моргания пользователя с использованием метрики EAR (Eye Aspect Ratio, рис. 5), которая позволяет точно и непрерывно отслеживать моменты моргания, фиксируя движения век.

Метрика EAR рассчитывается на основе расстояний между ключевыми точками вокруг глаза. Эти точки, обычно получаемые с помощью технологии распознавания лиц, как в нашем случае с помощью mediapipe, обозначают края верхнего и нижнего века. При моргании расстояние между веками сокращается, и метрика EAR падает. Когда глаз полностью закрывается, EAR приближается к нулю, что и используется для фиксации события моргания. Уравнение 4 описывает расчет этой метрики.

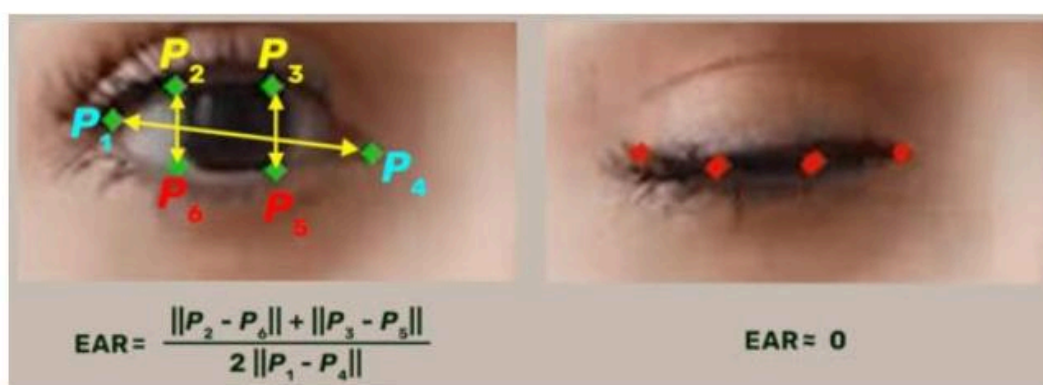


Рис. 5. EAR

Этот коэффициент стабилен, когда глаз открыт, и резко падает к нулю при закрытии глаза. Таким образом, EAR — это простая, но эффективная метрика, которая практически не меняется при открытом глазе и значительно снижается, когда пользователь моргает

$$EAR = \frac{||P_2 - P_6|| + ||P_3 - P_5||}{2 \cdot ||P_1 - P_4||}. \quad (4)$$

Запись и анализ данных о морганиях в перспективе позволит проводить более сложный анализ поведения пользователя. Вот несколько возможных применений:

Анализ концентрации. Моргания часто указывают на изменение концентрации и утомляемость. Долгие промежутки между морганиями могут свидетельствовать о высокой концентрации, тогда как частое моргание — признак усталости.

Эмоциональный анализ. Моргания, а также изменения в скорости и частоте морганий, могут быть индикаторами эмоциональных состояний, таких как стресс или волнение. Это может быть использовано в приложениях, которые анализируют реакции пользователя на контент.

Поддержка анализа пользовательских взаимодействий. Если пользователь активно взаимодействует с графическими объектами на экране, метрика EAR может фиксировать моменты, когда он, например, моргает после

переключения внимания на новый элемент интерфейса, что позволяет анализировать паттерны восприятия информации.

На рис. 6 приведены примеры визуализации взгляда пользователя при просмотре им различных объектов и изображений на экране.

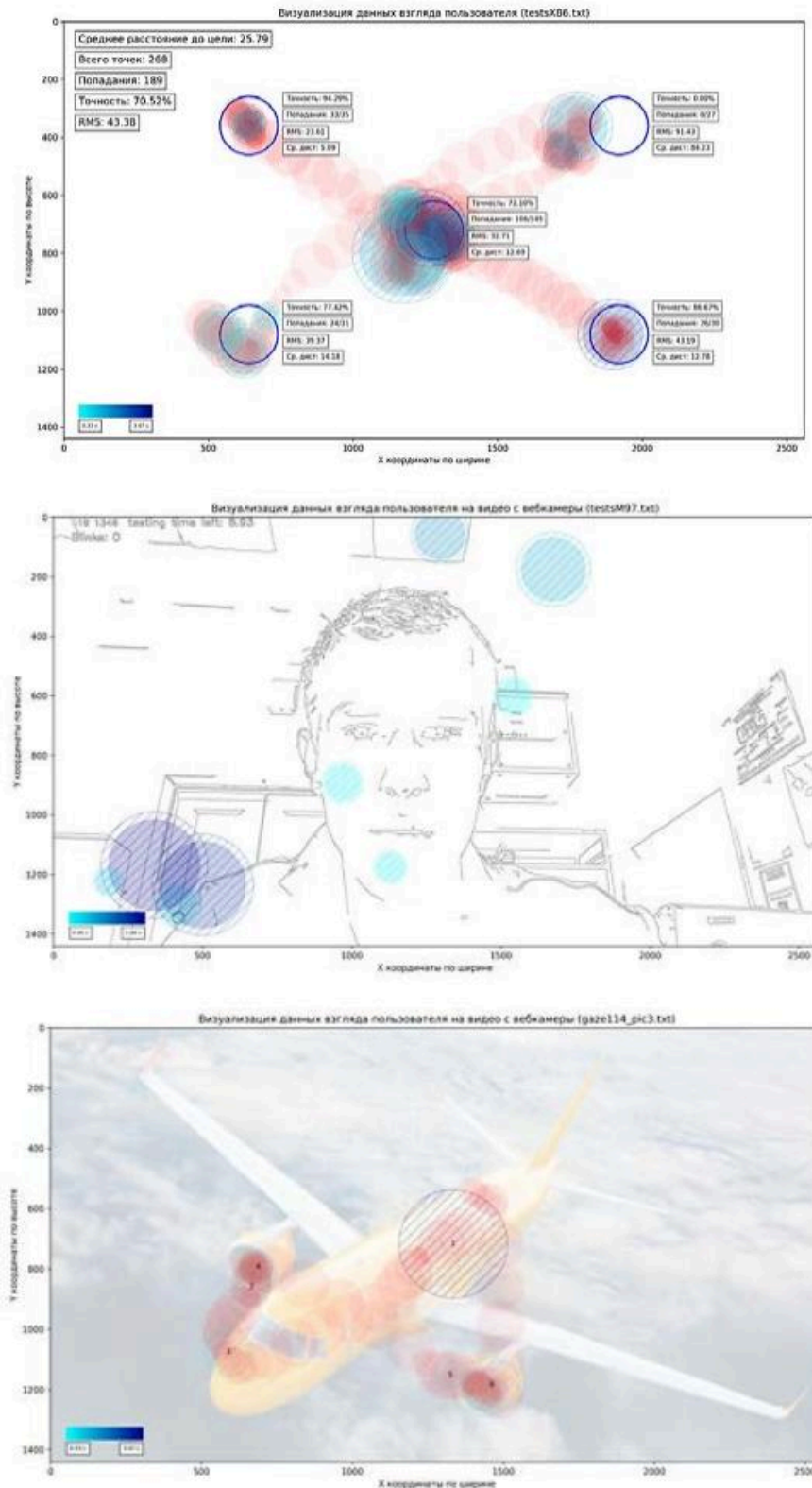


Рис. 6. Визуализация данных взгляда пользователя

Данные, которые регистрирует модернизированный трекер с расширенным функционалом, позволяют записывать паттерны поведения взгляда пользователя с различными графическими интерфейсами, анализировать различные объекты интереса, изучать акценты внимания пользователя и др.

Обсуждение полученных результатов. Диаграмма рассеивания в модернизированном трекере значительно улучшает понимание того, как именно перемещается взгляд пользователя. В отличие от обычной тепловой карты. Она позволяет более точно проследить траекторию движения глаз. Каждая точка на диаграмме представляет конкретное положение взгляда, что делает возможным анализ даже мелких и быстрых движений, которые важно учитывать для точной оценки.

Кроме того, эта диаграмма помогает увидеть не только точку фиксации взгляда, но и радиус зоны внимания, что дает более полное представление о том, как пользователь взаимодействует с объектами на экране. Важно понимать не только, куда человек смотрит, но и насколько точно он удерживает внимание в нужной области. Эти преимущества видны при сравнении диаграмм рассеивания и тепловых карты базового трекера [7] с подобными у модифицированной версии (рис. 7–9).

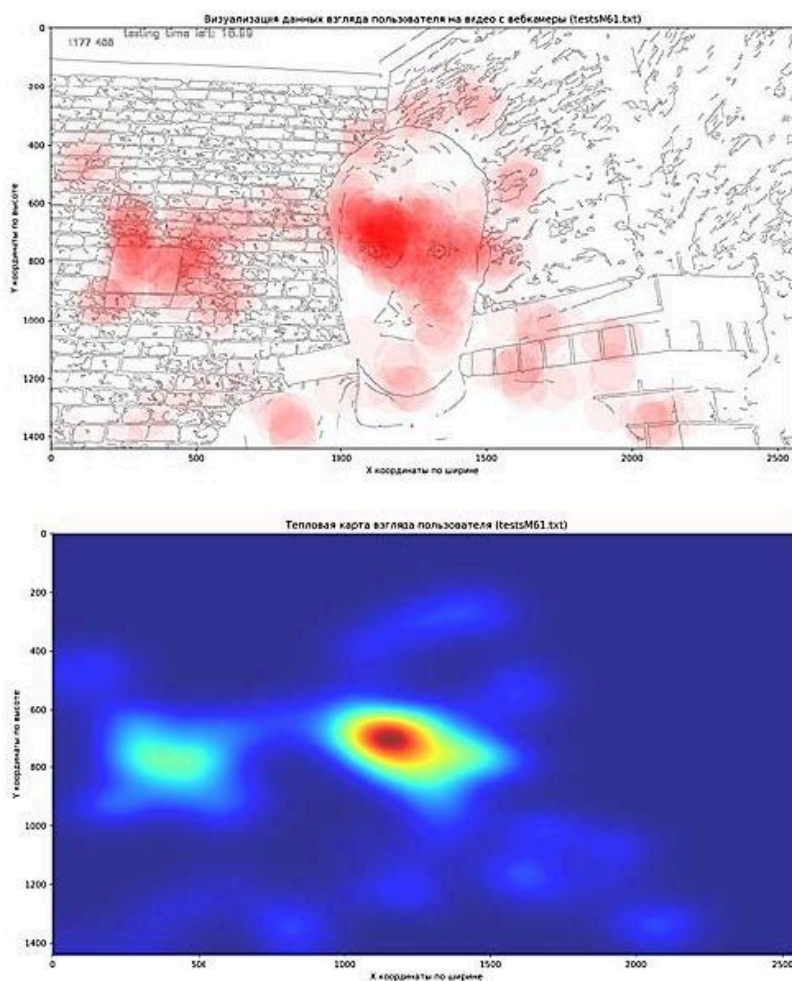


Рис. 7. Диаграмма разброса и тепловая карта базового трекера

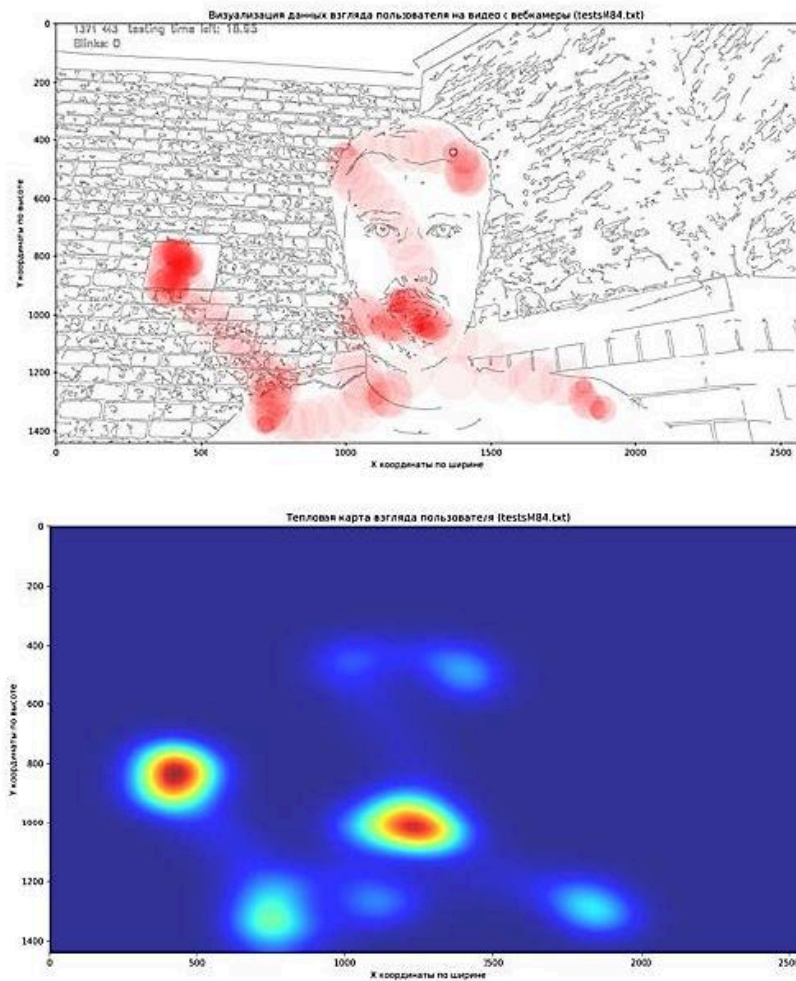


Рис. 8. Диаграмма разброса и тепловая карта модифицированного трекера

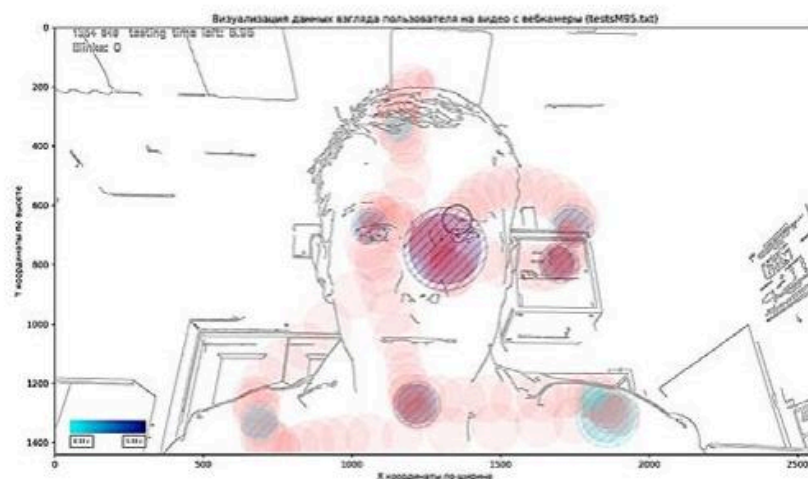


Рис. 9. Диаграмма разброса с фиксациями взгляда

Наложение полупрозрачных кругов на диаграмму также играет важную роль: по мере увеличения числа точек взгляда в одной и той же зоне, интенсивность цвета усиливается. Это визуально аналогично тепловой карте, но с дополнительной детализированной информацией. Чем ярче становится об-

ласть, тем дольше пользователь фокусировался на ней, что делает диаграмму мощным инструментом для анализа точек максимального интереса и концентрации внимания.

Визуализация данных взгляда совместно с зарегистрированными фиксациями позволяет не только дополнительно подтвердить то, что в зоне с высокой интенсивностью цвета пользователь фиксировал свой взгляд, но и то, сколько именно продолжалась фиксация и какой разброс взгляда был у пользователя, что говорит нам о том, как именно он изучал объект на экране.

Заключение. Проведенная работа продемонстрировала успешное расширение функционала самописного трекера глаз для повышения эффективности анализа когнитивного восприятия сложных графических изображений. В ходе исследования была внедрена метрика EAR для точного отслеживания морганий, реализован адаптивный указатель направления взгляда, а также разработан алгоритм регистрации и отслеживания фиксаций взгляда. Эти усовершенствования позволят увеличить качество анализа поведенческих и когнитивных паттернов пользователей при взаимодействии с визуальными элементами. Таким образом, разработанный трекер глаз представляет собой эффективный инструмент для глубокого изучения визуального восприятия, который может найти применение как в научных исследованиях, так и в практических задачах, связанных с улучшением интерфейсов и созданием адаптивных визуальных систем.

Список источников

- [1] Горячкин Б.С. *Восприятие информации в контуре управления человеко-машинных систем*. Москва, Изд-во «Спутник +», 2024, 376 с.
- [2] Савельев А.А., Горячкин Б.С. Построение эффективных сложных информационных моделей на основе трекинга глаз. *Технологии дополненной и виртуальной реальности. I науч.-практ. конф.: сб. ст.* Москва, Изд-во МГТУ им. Н.Э. Баумана, 2025, с. 175–185.
- [3] Мунипов В.М., Зинченко В.П. *Эргономика: человекоориентированное проектирование техники, программных средств и среды*. Москва, Логос, 2001, 356 с.
- [4] Горячкин Б.С. Эргономический сертификат автоматизированной системы обработки и отображения информации и управления. *Международный научно-исследовательский журнал*, 2016, № 9–2 (51), с. 25–29.
- [5] *Основы оптимального стиля. Когнитивно-ориентированный анализ удобочитаемости текста программы. Фовеа, поле зрения*. URL: https://habr.com/ru/companies/epam_systems/articles/517398/ (дата обращения 12.09.2025).
- [6] Горячкин Б.С., Савельев А.А., Лобанов Д.С., Фонин М.А. Классификация и сравнительный анализ методов машинного обучения для обработки и моделирования глазодвигательной активности. *Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление*, 2026, № 1, с. 160–169.
- [7] Горячкин Б.С., Якубов А.Р., Аникин Ф.А. Разработка инструментария трекинга глаз. *Нейрокомпьютеры: разработка, применение*, 2024, т. 26, № 2, с. 5–12.