

МЕТОД ОЦЕНКИ ОБЛАСТЕЙ ИНТЕРЕСА ПРИ АЙТРЕКИНГЕ
Коровин Я.С.¹, Хисамутдинов М.В.², Фисунов А.В.³, Иванов Д.Я.⁴
Email: Korovin695@scientifictext.ru

¹Коровин Яков Сергеевич – кандидат технических наук, директор;

²Хисамутдинов Максим Владимирович – кандидат технических наук, старший научный сотрудник;

³Фисунов Александр Владимирович – программист;

⁴Иванов Донат Яковлевич – кандидат технических наук, старший научный сотрудник,

Научно-исследовательский институт многопроцессорных вычислительных систем им. академика А.В. Каляева
Южный федеральный университет,
г. Таганрог

Аннотация: современные многомодальные интерфейсы взаимодействия оператора и управляемого объекта (робота, мехатронного комплекса и т.п.) могут использовать не только визуальный и акустический каналы связи, но и отслеживание направления взгляда оператора. Современные системы отслеживания направления глаз оператора зачастую используют дорогостоящее и сложное в эксплуатации оборудование. Однако существует спрос на дешевые и простые в применении системы. Один из подходов к созданию таких систем заключается в применении легкой конструкции очков оператора, в которой функции экрана, камеры, вычислительного и телекоммуникационного устройства выполняет смартфон бюджетного класса с установленным специальным программным обеспечением. В данной работе предложен метод оценки областей интереса при распознавании направления взгляда оператора в такой системе айтрекинга. Алгоритмическая и программная реализация предложенного метода позволяет применять его на доступных массовых устройствах.

Ключевые слова: отслеживание взгляда, интерфейс управления, айтрекинг, область интереса, распознавание образов.

METHOD FOR ASSESSING AREAS OF INTEREST IN EYE TRACKING
Korovin Ya.S.¹, Khisamutdinov M.V.², Fisunov A.V.³, Ivanov D.Ya.⁴

¹Korovin Yakov Sergeevich – PhD in Tecnology, Director;

²Khisamutdinov Maxim Vladimirovich – PhD in Tecnology, Senior Researcher;

³Fisunov Alexander Vladimirovich - Programmer;

⁴Ivanov Donat Yakovlevich – PhD in Tecnology, Senior Researcher,

SCIENTIFIC RESEARCH INSTITUTE OF MULTIPROCESSOR COMPUTER SYSTEMS NAMED AFTER ACADEMICIAN
A.V. KALYAEV
SOUTHERN FEDERAL UNIVERSITY,
TAGANROG

Abstract: modern multimodal interfaces of interaction between the operator and the controlled object (robot, mechatronic complex, etc.) can use not only visual and acoustic communication channels, but also tracking the direction of the operator's gaze. Modern operator eye tracking systems often use expensive and complex equipment. However, there is a demand for cheap and easy-to-use systems. One of the approaches to the creation of such systems is to use a lightweight design of operator glasses, in which the functions of a screen, camera, computing and telecommunication device are performed by a budget class smartphone with installed special software. This paper proposes a method for assessing areas of interest when recognizing the direction of the operator's gaze in such an eye tracking system. Algorithmic and software implementation of the proposed method makes it possible to apply it on available mass devices.

Keywords: eye tracking, control interface, eye tracking, region of interest, pattern recognition.

УДК 004.93

Введение

Айтрекинг – определение направления взгляда и его изменение во времени [1, 2] являются важными параметрами при оценке внимания, потребностей и эмоционального состояния человека [3].

В одной и той же ситуации, движение и геометрические характеристики глаз уникальны для каждого человека, что делает решение задач локализации зрачков, оценки направления взгляда важными для многих прикладных задач, таких как построение систем безопасности с использованием биометрических данных, оценка эмоционального состояния человека, анализ внимания или построение бесконтактного интерфейса взаимодействия человека с компьютером [4, 5] или роботом [6].

Для получения положения зрачка из видеопоследовательности и определения направления взгляда человека было разработано множество методов. Эти методы делятся на две основные группы [7, 8]: appearance-based [5, 8–14] и feature-based [9, 15]. Также существуют model-based решения, предложенные

в работах [2, 4, 15–17]. Существуют также гибридные алгоритмы [15, 18], которые совмещают в себе feature- и model-based решения. К этой группе относятся: алгоритм Starburst [15], выполняющий поиск набора точек, описывающий контур зрачка с помощью RANSAC метода [17].

Множество работ посвящено отслеживанию направления взгляда и оценке областей интереса на экране монитора при использовании специализированного, и зачастую дорогостоящего, оборудования. Подобного рода системы могут включать в себя набор камер, расположенных под монитором, или специализированные очки с миниатюрными камерами и ИК-подсветкой [2, 4]. Наиболее оптимальным решением является использование инфракрасного источника света с длиной волны около 880 н.м., которая невидима для глаза и может быть захвачена промышленными камерами [3]. Такие системы дороги в производстве и сложны в эксплуатации. Актуальной [16, 19] является задача разработки дешевых систем, позволяющих применять их в различных областях, таких как дистанционное управление робототехническими комплексами и другими мехатронными устройствами.

Предлагаемый подход

Предлагается в качестве устройства для отображения графической информации очков оператора, получения видеопотока с изображением глаз человека, и обработки этого видеопотока с целью определения областей интереса, использовать смартфон на платформе Android со специализированным программным обеспечением.

На экране смартфона может отображаться графическая информация, фронтальная камера смартфона может захватывать область глаз человека, а установленное ПО обрабатывать получаемые данные (перемещение глаз во время просмотра видеоролика), сохранять их, и передавать на сервер для хранения. Для этого предлагается использовать специальные очки (рисунок 1), описанные в работе [6], в которые может быть установлен смартфон на платформе Android, и программное обеспечение, реализующее предложенный в данной работе метод оценки областей интереса.

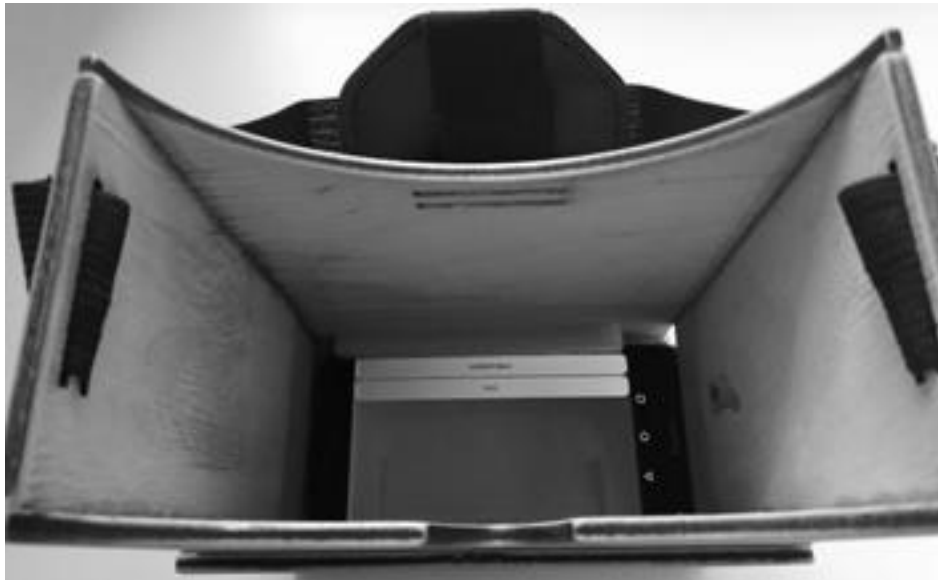


Рис. 1. Прототип очков с креплением под телефон, и ремнями для фиксации на голове [6]

Предлагаемый метод включает в себя калибровку системы, детектирование центра зрачка, определение области интереса, сохранение и передачу полученных данных.

Детектирование зрачка и локализация центра

Для определения центра зрачка используется метод, описанный в работе [6]. Основой данного метода является предположение о том, что направления векторов, построенных из центра зрачка к каждой точке изображения глаза, должны совпадать с направлением градиентом яркости в указанной точке изображения. Следовательно, центр зрачка можно найти путем поиска максимума функции, представленной в выражении 1.

$$c^* = \arg \max \left\{ \frac{1}{N} \sum_{i=1}^N (d_i^T g_i)^2 \right\} \quad (1)$$

где c^* – возможный центр зрачка;
 d_i – вектор смещения из точки p_i в c^* ;
 g_i – градиент изображения в точке p_i .

Перед вычислениями область с изображением глаз преобразуется в изображение в градациях серого. Затем метод перебирает все пиксели на изображении глаза и принимает решение о координатах центра

зрачка на основе максимума функции 1. На рисунке 2 схематично показано влияние координат точки c на вычисление функции (1).

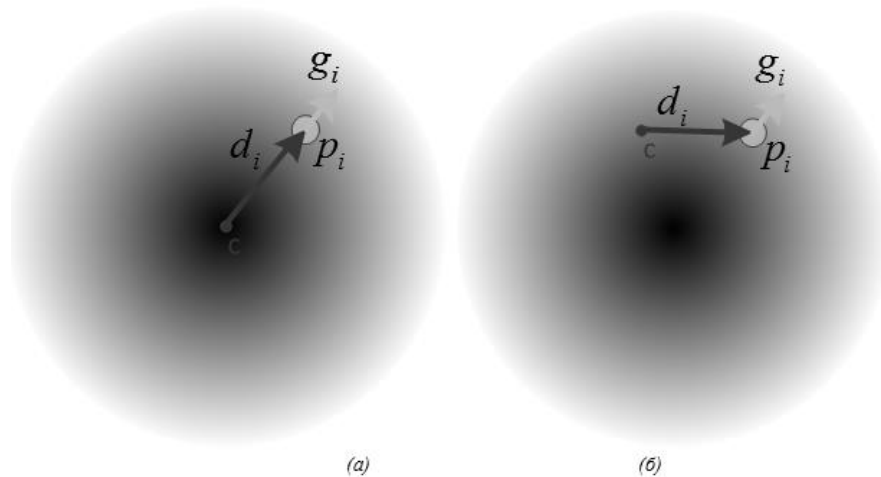


Рис. 2. Влияние координат точки c на вычисление функции: (а) направления вектора, построенного из центра зрачка c к точке p_i в области радужки, и вектора градиента в точке p_i совпадают, (б) – для точки c не являющейся центром, направления совпадать не будут, следовательно, дадут меньший вклад в функцию (1)

Зачастую на изображении глаза наблюдается не такое идеальное распределение яркостей, как показано на рисунке 2, например, веки и ресницы в области радужной оболочки глаза, общая низкая контрастность изображения, и блики в области зрачка, могут вносить погрешность в расчет функции (1), следовательно, снижать точность определения центра зрачка. Для компенсации подобного негативного влияния добавляется весовой коэффициент w_c :

$$\frac{1}{N} \sum_{i=1}^N w_c (d_i^T g_i)^2 \quad (2)$$

$$w_c = I^*(cx, cy) \quad (3)$$

где: w_c - это значение серого в точке (cx, cy) сглаженного и перевернутого входного изображения I^* .

Для компенсации бликов в области зрачка последовательно применяется два морфологических оператора: дилатация и эрозия, с ядром размера 7×7 пикселей.

Результатом работы, описанного выше метода, будут являться координаты центра зрачка на изображении, получаемого с фронтальной камеры смартфона. На следующих этапах, эти координаты необходимо связать с координатами точки интереса на экране смартфона.

Калибровка

Для сопоставления координат центра зрачка с координатами экрана требуется выполнить калибровку системы, которая позволит подстроить работу под индивидуальные особенности глаз конкретного человека.

Процесс калибровки подобных систем обычно состоит в том, что пользователь поочередно сосредоточивает внимание на последовательном наборе точек на экране. Стандартный набор для калибровки состоит из 25 точек, равномерно распределенных по экрану. На рисунке 3 показан пример калибровочного набора.

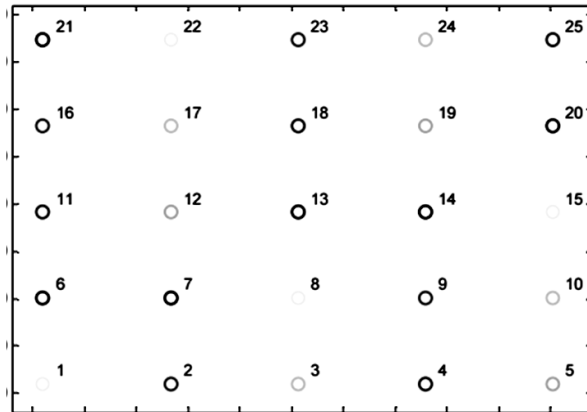


Рис. 3. Набор точек для калибровки системы отслеживания взгляда

Однако, для экрана смартфона, находящегося на расстоянии нескольких десятков сантиметров от глаз пользователя, такое количество точек является избыточным. В результате тестов различных калибровочных паттернов, было принято решение выполнять калибровку с помощью последовательно показываемых 4 точек, располагающихся по углам изображения. Уменьшение количества точек позволило упростить и ускорить процесс калибровки, не снижая уровень точности. Итоговый набор точек калибровки показан на рисунке 4.



Рис. 4. Калибровочный набор для системы оценки внимания пользователя на экране смартфона

На этапе калибровки пользователю необходимо переключать точки и указывать системе, что он сосредоточил взгляд в показываемой точке. Поскольку телефон находится в этот момент в корпусе очков, то для управления процессом калибровки используется bluetooth джойстик. После калибровки данные записываются в локальное хранилище приложения и используются в дальнейшей работе системы.

На рисунке 5 представлен экран калибровки, на котором видно, что поиск центра зрачка осуществляется в заранее выбранной области (синий прямоугольник). Такое допустимо из-за того, что при надевании очков область глаза находится в одном и том же месте изображения. При необходимости эту область можно переместить в настройках приложения.

Окружность в верхнем левом углу представляет собой маркер, на котором пользователь должен сосредоточить взгляд. Белая рамка на изображении служит для корректировки яркости изображения, поскольку телефон находится в корпусе очков, а показываемое видео может быть темным, то необходима дополнительная подсветка, иначе зрачок на изображении не будет виден.

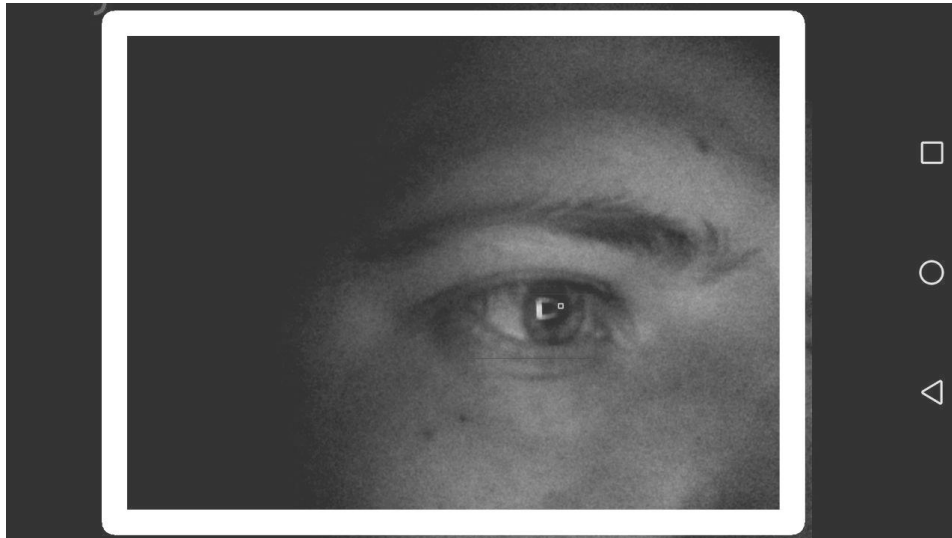


Рис. 5. Экран калибровки приложения

После выполнения калибровки необходимо провести пересчет координат зрачка в координаты точки на экране, которой они соответствуют.

Преобразование координат центров глаз в координаты экрана

Если предположить, что экран смартфона имеет разрешение 1920x1080, то мы знаем координаты каждого из углов экрана: (0, 0), (1920, 0), (1920, 1080) и (0, 1080). Обозначим эти точки как: (sx0, sy0), (sx1, sy0), (sx1, sy1), (sx0, sy1). Во время калибровки, когда глаз смотрит на каждый из углов экрана, координаты центра зрачка также образуют квадрат, давайте по тому же принципу обозначим координаты углов этого малого квадрата: (ex0, ey0), (ex1, ey0), (ex1, ey1), (ex0, ey1).

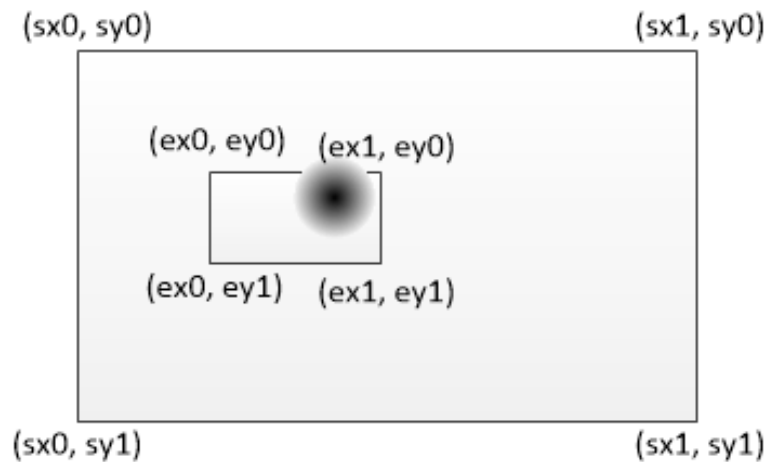


Рис. 6. Пересчет координат

Зная эти координаты, и координаты центра зрачка (ix, iy), можно получить координаты точки интереса (px, py), которые вычисляются следующим образом:

$$px = (ix - ex0)/(ex1 - ex0) * sx1; \quad (4)$$

$$py = (iy - ey0)/(ey1 - ey0) * sy1; \quad (5)$$

В результате вычислений будут получены координаты точки интереса пользователя на экране смартфона. Далее необходимо накапливать эту информацию для дальнейшего сохранения и анализа.

Определение направления взгляда

В качестве структуры данных для хранения информации об областях интереса выступает heatmap, которая представляет собой двумерный целочисленный массив, размер которого совпадает с разрешением экрана смартфона. На каждом кадре, для каждой точки интереса, в радиусе 5% от длины и ширины экрана происходит накопление значений этого массива, при этом, чем ближе точка к центру области интереса, тем больший вес ей присваивается. Все элементы массива, не входящие в эту окружность, уменьшают свое значение на 1. Тем самым наиболее высокое значение будет иметь область,

на которую смотрит пользователь длительное время, а области, на которые он смотрел ранее, постепенно сбросят значение до нуля.

Оценка точности определения области интереса

Для оценки точности работы метода был проведен следующий эксперимент. Десяти пользователям был показан набор из 20 точек, равномерно распределенных по экрану, точки появлялись последовательно по нажатию на кнопку bluetooth джойстика.

Задачей пользователей было сосредоточить внимание на показываемой точке и нажать на кнопку, в результате чего рассчитывалось расстояние между координатами показываемой точки шаблона, и точки интереса, определяемой программой. При этом все координаты переводились из абсолютных в относительные.

Полученные расстояния для всех пользователей и всех точек были просуммированы и усреднены, что в результате дало относительную погрешность измерения координат точки интереса, которая составила 12.4%. Основной вклад в погрешность вносит центральная часть экрана, так как при взгляде вперед, когда человек смотрит на соседние точки на экране, зрачок может перемещаться очень незначительно, а следовательно, метод не сможет определить, на какую именно точку смотрит глаз (при условии, что эти точки расположены достаточно близко).

Выводы

В работе предложен метод оценки областей интереса на экране смартфона. Относительная погрешность определения точки интереса на экране смартфона составляет 12.4%.

Благодарность: Работа выполнена при финансовой поддержке РФФИ проект № 18-29-03225.

Список литературы / References

1. Фроимсон М.И. и др. Система определения направления взгляда пользователя в режиме реального времени // Спецтехника и связь. Общество с ограниченной ответственностью «Спецтехника и связь», 2013. № 3. С. 32–34.
2. Hansen D.W., Ji Q. In the eye of the beholder: A survey of models for eyes and gaze // IEEE Trans. Pattern Anal. Mach. Intell. IEEE, 2009. Vol. 32. № 3. P. 478–500.
3. Morimoto C.H., Mimica M.R.M. Eye gaze tracking techniques for interactive applications // Comput. Vis. image Underst. Elsevier, 2005. Vol. 98. № 1. P. 4–24.
4. Lupu R.G., Ungureanu F., Siriteanu V. Eye tracking mouse for human computer interaction // 2013 E-Health and Bioengineering Conference (EHB), 2013. P. 1–4.
5. Sugano Y. et al. Appearance-based gaze estimation with online calibration from mouse operations // IEEE Trans. Human-Machine Syst. IEEE, 2015. Vol. 45. № 6. P. 750–760.
6. Хисамутдинов М.В., Коровин Я.С., Иванов Д.Я. Расширенный интерфейс человек-робот с дополненной реальностью // Вестник Белгородского государственного технологического университета им. ВГ Шухова, 2019. № 9. С. 113–120.
7. Chennamma H.R., Yuan X. A survey on eye-gaze tracking techniques // arXiv Prepr. arXiv1312.6410. 2013. P. 388–393.
8. Lu F. et al. A head pose-free approach for appearance-based gaze estimation. // BMVC, 2011. P. 1–11.
9. Wang Y. et al. Appearance-based gaze estimation using deep features and random forest regression // Knowledge-Based Syst. Elsevier, 2016. Vol. 110. P. 293–301.
10. Lu F. et al. Inferring human gaze from appearance via adaptive linear regression // 2011 International Conference on Computer Vision, 2011. P. 153–160.
11. Lai C.-C. et al. Appearance-based gaze tracking with free head movement // 2014 22nd International Conference on Pattern Recognition, 2014. P. 1869–1873.
12. Hansen D.W. et al. Eye typing using Markov and active appearance models // Sixth IEEE Workshop on Applications of Computer Vision, 2002. (WACV 2002). Proceedings, 2002. P. 132–136.
13. Ishikawa T. Passive driver gaze tracking with active appearance models // Proceedings of the 11 World Congress on Intelligent Transportation Systems. figshare, 2004.
14. Tan K.-H., Kriegman D.J., Ahuja N. Appearance-based eye gaze estimation // Sixth IEEE Workshop on Applications of Computer Vision, 2002. (WACV 2002). Proceedings, 2002. P. 191–195.
15. Li D., Winfield D., Parkhurst D.J. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops, 2005. P. 79–87.
16. Ferhat O., Vilariño F. Low cost eye tracking // Comput. Intell. Neurosci. Hindawi Publishing Corp., 2016. Vol. 2016. P. 17.
17. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // Commun. ACM. ACM, 1981. Vol. 24. № 6. P. 381–395.

18. *Nguyen P. et al.* Calibration-free gaze tracking using particle filter // 2013 IEEE International Conference on Multimedia and Expo (ICME), 2013. P. 1–6.
19. *Малин И.* Отслеживание направления взгляда в реальном времени без использования специальной видеоаппаратуры // Труды «23-й международной конференции по компьютерной графике и зрению». Институт автоматизации и процессов управления ДВО РАН, Дальневосточный федеральный университет, 2013. Ч. 294–297.