

Технология отслеживания взгляда в системах дополненной реальности

Ю.А. Беляева, А.Д. Луговой, В.А. Селиванова

Московский авиационный институт (Национальный исследовательский университет), г. Москва

Аннотация: В статье рассматриваются принципы работы, ключевые технологии и перспективы развития систем отслеживания взгляда в устройствах виртуальной реальности. Освещены основные компоненты таких систем, включая инфракрасные камеры, алгоритмы компьютерного зрения и методы калибровки. Подробно анализируются технологии отслеживания взгляда, такие как роговичное отражение в центре зрачка (Pupil Center Corneal Reflection - PCCR), а также их интеграция с рендерингом для реализации фовеального рендеринга, что позволяет значительно снизить нагрузку на GPU. Рассмотрены актуальные проблемы, включая задержки и энергопотребление, и предложены пути их решения, такие как использование предиктивных алгоритмов и аппаратного ускорения. Особое внимание уделено перспективным направлениям, включая нейроинтерфейсы и голографические системы. Статья основана на последних исследованиях и разработках ведущих компаний. Материалы статьи представляют интерес для разработчиков устройств виртуальной реальности, исследователей в области человеко-машинного взаимодействия и специалистов по компьютерному зрению.

Ключевые слова: отслеживание взгляда, виртуальная реальность, фовеальный рендеринг, компьютерное зрение, человеко-машинное взаимодействие

Введение

Современные технологии виртуальной и дополненной реальности (VR/DR) стремительно развиваются, расширяя границы взаимодействия человека с цифровыми средами. Одним из ключевых направлений, способствующих повышению реалистичности и эффективности таких систем, является технология отслеживания взгляда. Эта технология позволяет точно определять, куда направлен взгляд пользователя, и использовать полученные данные для адаптации визуального контента, управления интерфейсом и анализа поведения в реальном времени. В контексте VR -технологий отслеживания взгляда играет всё более значимую роль. Она открывает возможности для реализации фовеального рендеринга, повышения интерактивности, адаптивного управления визуальными эффектами и улучшения пользовательского опыта за счёт снижения

когнитивной нагрузки. Наряду с этим, отслеживание взгляда используется для научных исследований, требующих оценки уровня вовлечённости и внимания, охватывая широкий спектр вопросов, начиная от авиапромышленности и подготовки летного состава, медиасферы, приборостроения, медицины и заканчивая нейромаркетингом и исследованием потребительского спроса, что позволяет доказать актуальность данной технологии в современном информационном пространстве.

Несмотря на высокие перспективы, технология отслеживания взгляда в ВР/ДР сталкивается с рядом технических и инженерных вызовов. Среди них – обеспечение высокой точности и надёжности слежения, минимизация задержек, необходимость калибровки, а также интеграция с аппаратной частью шлемов виртуальной реальности. Настоящая работа посвящена анализу современных методов реализации отслеживания взгляда в системах виртуальной реальности, оценке их возможностей и ограничений, а также обзору направлений, в которых ведутся наиболее активные исследования и разработки.

Цель исследования – обобщить исследования, посвященные технологии отслеживания взгляда, а также рассмотреть возможные способы применения данной технологии в дополненной реальности.

История вопроса. Технология отслеживания взгляда объектом пристального внимания ученых стала относительно недавно, но она активно применяется для решения широкого спектра задач как в отечественном, так и в международном научном сообществе. Наиболее распространенной сферой применения технологии стали визуальные системы; психология и нейропсихология; психолингвистика и здоровье-сбережение; маркетинг и сфера PR; экономика; программное обеспечение; системы слежения, обеспечения безопасности воздушных судов и другого вида транспорта;

создание виртуальной реальности. В современном информационном пространстве трудно найти область, которая не подверглась обработке рассматриваемой нами технологии, поскольку она активно проникает во все сферы жизнедеятельности человека, облегчая обработку данных в технических и в гуманитарных отраслях знания, неся огромный исследовательский потенциал и обеспечивает актуальность проведенного нами исследования.

Методы исследования. Нами активно используются теоретические методы исследования: описательный, метод сплошной выборки, что позволяет рассмотреть текущее состояние изучения технологии отслеживания взгляда, а также экспериментальные методы, позволяющие оценить степень влияния технологии в современных системах виртуальной реальности.

Технология отслеживания взгляда представляет собой важный элемент современных систем виртуальной реальности, благодаря которой можно определить точное направление взгляда пользователя и выйти на новый уровень взаимодействия человека с виртуальными объектами, позволяя глубоко проникнуть в мир дополненной реальности.

1. Типы движений глаз

Существует несколько различных типов движений глаз, которые выполняют разные функции и контролируются разными нейронными цепями. Оптимизация функции отслеживания взгляда для приложений виртуальной реальности часто требует идентификации и их различения, а также использования знаний о визуальной и когнитивной обработке во время зрительного контакта для достижения желаемого результата [1].

Необходимость движений глаз возникает из-за того, что острота зрения человека не является одинаковой во всем поле зрения. Острота зрения

максимальна в фовеа, центральной области сетчатки [2]. В области фовеа находится наибольшая плотность фоторецепторов, что позволяет видеть части поля зрения, попадающие на нее, с максимальной детализацией [3]. За пределами области фовеа острота зрения постепенно снижается к краям сетчатки.

Субъективное ощущение возможности видеть все поле зрения с высокой остротой достигается путем сканирования поля зрения последовательными движениями глаз, известными как саккады, и интеграции полученных снимков сцены с высокой остротой, процесс, известный как транс-саккадическая интеграция [4]. Скорость саккад может достигать 400–600 градусов в секунду, их продолжительность составляет от 30 до 120 миллисекунд, а амплитуда — от 1 до 45 градусов; между амплитудой, продолжительностью и скоростью существует систематическая зависимость, известная как основная последовательность саккад [5]. Движения глаз при слежении происходят, когда глаза отслеживают медленно движущуюся визуальную цель, чтобы стабилизировать изображение этой цели на сетчатке.

Другой тип движения глаз, вестибуло-окулярный рефлекс (VOR). Эти движения приводят к повороту глаз, чтобы свести к нулю движение изображения на сетчатке, которое возникает в результате движения головы относительно неподвижной окружающей среды. Рефлекс в основном управляется вестибулярными сигналами, которые преобразуют инерционные сигналы в движение головы, как линейное, так и угловое [6]. Результирующее стабилизирующее движение глаз, называемое в совокупности визуально усиленным вестибуло-окулярным рефлексом, позволяет нам видеть мир четким и стабильным, несмотря на движение головы.

Неточная и/или неточная классификация и количественная оценка, в свою очередь, ставит под угрозу использование данных о движениях глаз для

приложений виртуальной реальности. Поэтому передовые методы идентификации и классификации, использующие всю доступную информацию, являются активной областью исследований [7].

2. Методы отслеживания глаз в VR HMD

Для отслеживания движений глаз в HMD использовались три метода: электроокулография (ЭОГ), склеральные поисковые катушки и наиболее распространенная, видеоокулография (ВОГ). ЭОГ измеряет ориентацию глаза путем размещения электродов на коже вокруг глаза, которые измеряют потенциал покоя глаза. Электроды могут быть легко встроены в устройство в месте соприкосновения с лицом [8]. Этот метод работает, потому что глаз является диполем, который заряжен положительно в направлении роговицы и отрицательно в направлении сетчатки. Разница в напряжении между электродами, размещенными на противоположных сторонах глаза, отображает ориентацию глаза [9]. ЭОГ — единственный метод, который позволяет отслеживать движение глаз, когда они закрыты.

Метод скеральной поисковой катушки работает путем отслеживания ориентации проволоочной петли, встроенной в контактную линзу, которую носит пользователь [10]. Голова пользователя располагается между катушками Гельмгольца, которые генерируют однородное магнитное поле. Когда глаза движутся в однородном и известном магнитном поле, в скеральных катушках индуцируется электрический ток, указывающий на горизонтальную, вертикальную и торсионную ориентацию глаза. Этот метод является высокоточным и точным, с заявленным пространственным разрешением менее 0,1 и временным разрешением более 1 кГц [11].

На сегодняшний день наиболее распространенным методом отслеживания движения глаз, используемым в устройствах виртуальной реальности, является ВОГ [12]. В настоящее время коммерчески доступные устройства отслеживания движения глаз, включая Tobii, Pupil Labs, Varjo и

Fove, используют отслеживание движения глаз на основе ВОГ. Изображения глаз захватываются камерами, установленными в устройстве, и анализ видеокадров указывает на ориентацию глаз.

3. Технические аспекты отслеживания движения глаз

3.1. Качество данных отслеживания движения глаз

Высокое качество данных слежения за движением глаз является важным условием для любого приложения, использующего данные о движении глаз. Качество данных слежения за движением глаз относится к производительности устройства слежения за движением глаз, измеряемой с помощью различных показателей, включая пространственную точность, пространственную точность, частоту дискретизации и задержку.

Пространственная точность — это среднее угловое отклонение между истинным положением взгляда и положением взгляда, оцененным устройством слежения за движением глаз [13]. Она измеряется как расстояние в градусах между положениями целевой точки и средним положением набора данных, собранных устройством заданного временного интервала. Для точного измерения движений глаз с малой амплитудой, таких как фиксационные движения глаз и саккады с низкой амплитудой, требуются устройства слежения за движением глаз с высокой частотой дискретизации [14].

3.2. Калибровка

Калибровка — это процесс, с помощью которого система слежения за движением глаз находит функцию отображения, которая сопоставляет координаты, сообщаемые устройством слежения за движением глаз, с координатами точки взгляда в визуальной среде [15]. Большинство устройств слежения за движением глаз используют стандартную процедуру калибровки на основе точек. В ходе этой процедуры пользователю последовательно

показывается набор (обычно от 5 до 16) небольших точечных мишеней, и пользователю предлагается зафиксировать взгляд на каждой мишени в течение нескольких секунд. Функция отображения требует, чтобы пользователь фиксировал взгляд на каждой мишени в течение всего времени, пока мишень отображается в окружающей среде. В некоторых случаях, таких как взаимодействие на основе плавного слежения, используется только относительное движение глаз пользователя в реакции на объекты сцены, поэтому калибровка системы слежения за движением глаз не требуется [16].

3.3. Отображение и рендеринг

Создание по-настоящему иммерсивных опытов потребует разработки устройств, способных рендерить контент с визуальным качеством, близким к качеству человеческой зрительной системы. Для реализации таких систем отображения, по оценкам, необходимо будет передавать в устройство более 100 Гб/с данных. Человеческий глаз имеет очень высокую остроту зрения в центральной области сетчатки площадью 5,2% области сетчатки — фовеа. За пределами этой области острота зрения снижается по мере приближения к периферии сетчатки. Фовеа занимает около 4 % пикселей в потребительских системах виртуальной реальности [17], при этом около 96 % пикселей, отображаемых на дисплее, попадают в области сетчатки с низкой остротой зрения. Системы отображения, зависящие от взгляда, используют это явление для повышения производительности и снижения вычислительных затрат в устройстве, отображая контент, который находится за пределами фовеи, с низким разрешением или уровнем детализации [18].

3.4. Повышение эффективности рендеринга

Фовеальный рендеринг отображает области дисплея, которые находятся в центре поля зрения пользователя, с максимальным разрешением и снижает разрешение с увеличением эксцентриситета. Это приводит к

улучшению производительности и качества рендеринга и, как было показано, позволяет достичь экономии производительности до 50-70%.

Однако сжатие периферийных частей сцены может привести к появлению различных заметных артефактов, таких как туннельное зрение, наложение и мерцание, которые могут отвлекать пользователей и снижать степень погружения. Для устранения этих артефактов использовались три концентрических круга, центрированных на взгляде, с разрешением, постепенно ухудшающимся к периферии.

3.5. Уменьшение дискомфорта, вызванного конфликтами глубины

Обычные стереоскопические 3D-дисплеи, расположенные близко к глазам, такие как те, которые используются в ВР-шлемах, создают ощущение 3D, показывая каждому глазу отдельное 2D-изображение, причем каждое изображение отображается с небольшими различиями, чтобы создать бинокулярную диспаратет. Бинокулярная диспаратет является важным стимулом для вергенции, которая является важным сигналом глубины. Однако расстояние между глазами пользователя и изображением фиксировано расположением экрана дисплея. В результате, хотя 3D-изображения отображаются на различной глубине, глаза всегда фокусируются на одной глубине. Таким образом, дисплей не отображает правильное размытие сетчатки, что приводит к неспособности глаз правильно фокусироваться или аккомодироваться, вызывая потерю аккомодации — еще один важный ориентир глубины. Это несоответствие называется конфликтом вергенции и аккомодации и является источником визуального дискомфорта, включая напряжение глаз, помутнение зрения и головные боли [18].

Были предложены дисплеи с переменным фокусным расстоянием, дисплеи с монозрением, дисплеи с неизменной аккомодацией, мультифокальные дисплеи, дисплеи со сканированием сетчатки, дисплеи со

световым полем и голографические дисплеи. Из этих подходов дисплеи с переменным фокусом используют отслеживание движения глаз для смягчения этого конфликта. Дисплеи с переменным фокусом используют отслеживание движения глаз для активного отслеживания конвергенции глаз и используют фокусирующий элемент с переменным фокусом, чтобы соответствовать конвергенции глаз.

3.6. Уменьшение VR-тошноты

Также был исследован потенциал дисплеев, зависящих от взгляда, в уменьшении частоты возникновения VR-тошноты. Была исследована полезность ограничения поля зрения для уменьшения тошноты при использовании виртуальной реальности. Ограничение поля зрения (туннелирование) — техника для уменьшения визуально индуцированной тошноты, которая заключается в блокировании периферического зрения пользователей с помощью ограничителя, чтобы минимизировать оптический поток в периферических частях сетчатки, которые чувствительны к нему. Большинство современных реализаций этой технологии не реагируют на взгляд и могут снижать погружение и ощущение присутствия. Однако ограничение поля зрения с фокусировкой на центральной области зрения реализует ограничитель, который движется вместе с взглядом пользователя, что позволяет пользователям видеть большую часть визуальной сцены, при этом по-прежнему блокируя их периферическое зрение.

Заключение

Существующие технологии отображения продемонстрировали большой потенциал для повышения эффективности рендеринга и улучшения присутствия и комфорта в виртуальной реальности. Предыдущие исследования показали, что неиммерсивные дисплеи, зависящие от взгляда, негативно влияют на выполнение задач (например, скорость чтения). Кроме

того, большинство устройств отслеживания взгляда, интегрированных в системы виртуальной реальности, имеют высокую задержку, недостаточную точность и не имеют простых процедур калибровки. Задержка может оказать негативное влияние на рендеринг, зависящий от взгляда, вводя заметные задержки в конвейер рендеринга, что может снизить иммерсивность и вызвать дискомфорт в виртуальной реальности. Для решения проблемы задержки были предложены различные подходы, включая прогнозирование конечной точки саккады и положения взгляда на основе машинного обучения. Технологии отслеживания взгляда в системах виртуальной реальности не только улучшают взаимодействие пользователя с окружением, но и меняют сам подход к проектированию и созданию контента, что позволит комбинировать физическое восприятие с виртуальными впечатлениями.

Литература (References)

1. Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, van de Weijer J., Eye Tracking: a comprehensive guide to methods and measures. 2011, OUP Oxford. URL: books.google.com/books?id=5rIDPV1EoLUC
2. Cowey A, Rolls E, Human cortical magnification factor and its relation to visual acuity, 1974, Exp Brain Res 21(5). pp. 447–454. URL: doi.org/10.1007/BF00237163
3. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE Human photoreceptor topography, 1990, J Comparative Neurol 292(4). pp. 497–523. URL: doi.org/10.1002/cne.902920402
4. Melcher D, Colby CL Trans-saccadic perception, 2008 Trends Cognitive Sci 12(12). pp. 466–473. URL: doi.org/10.1016/j.tics.2008.09.003
5. Leigh RJ, Zee DS The Neurology of Eye Movements, 2015 OUP USA. URL: doi.org/10.1093/med/9780199969289.001.0001

6. Ramaioli C, Cuturi LF, Ramat S, Lehnen N, MacNeilage PR Vestibulo-ocular responses and dynamic visual acuity during horizontal rotation and translation, 2019, Front Neurol 10. P. 321. URL: doi.org/10.3389/fneur.2019.00321
 7. Toates FM Vergence eye movements, 1974, Documenta Ophthalmologica 37(1). pp. 153–214.
 8. Kothari R, Yang Z, Kanan C, Bailey R, Pelz JB, Diaz GJ Gaze-in-wild: a dataset for studying eye and head coordination in everyday activities, 2020, Scientific Rep 10(1) . pp. 1–18. URL: doi.org/10.1038/s41598-020-59251-5
 9. Bulling A, Roggen D, Tröster G Wearable EOG goggles: seamless sensing and context-awareness in everyday environments, 2009, J Ambient Intell Smart Environ 1(2). pp. 157–17112. Mowrer et al. 1935
 10. Robinson DA A method of measuring eye movement using a scleral search coil in a magnetic field, 1963, IEEE Tran Bio-Med Electron 10(4). pp. 137–145. URL: doi.org/10.1109/TBMEL.1963.4322822
 11. Whitmire E, Trutoiu L, Cavin R, Perek D, Scally B, Phillips J, Patel S EyeContact: scleral coil eye tracking for virtual reality, 2016, In: international symposium on wearable computers, digest of papers. URL: doi.org/10.1145/2971763.2971771
 12. Duchowski A Eye tracking techniques, 2007, In: eye tracking methodology: theory and practice, Springer London, London, pp. 51–59. URL: doi.org/10.1007/978-1-84628-609-4_5
 13. Holmqvist K, Nyström M, Mulvey F, Eye tracker data quality: what it is and how to measure it, 2012 In: eye tracking research and applications symposium (ETRA) 1(212). pp. 45–52. URL: doi.org/10.1145/2168556.2168563
 14. Richter C, Fromm CA, Diaz GJ, Hardware modification for improved eye tracking with the pupil labs virtual-reality integration, 2019, J Vis 19(10). pp. 147a–147a18. Shannon.
-

15. Harezlak K, Kasproski P, Stasch M, Towards accurate eye tracker calibration -methods and procedures, 2014, Procedia Comput Sci 35(2). pp. 1073–1081, URL: doi.org/10.1016/j.procs.2014.08.194
16. Zeng Z, Siebert FW, Venjakob AC, Roetting M, Calibration-free gaze interfaces based on linear smooth pursuit, 2020, J Eye Mov Res 13(1), URL: doi.org/10.16910/jemr.13.1.3
17. Bastani B, Turner E, Vieri C, Jiang H, Funt B, Balram N, Schade O, Foveated pipeline for AR/VR head-mounted displays, 2017, Inf Dis 33(6). pp. 14–35.
18. Patney A, Salvi M, Kim J, Kaplanyan A, Wyman C, Benty N, Luebke D, Lefohn A, Towards, Foveated rendering for gaze-tracked virtual reality, 2016, ACM Trans Gr 35(6):1–12. URL: doi.org/10.1145/2980179.2980246

Дата поступления: 24.06.2025

Дата публикации: 25.09.2025