



**ПРОФЕССИОНАЛЬНОЕ
ЗАВТРА**



VI Всероссийский сетевой конкурс студенческих проектов с участием студентов с инвалидностью

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФГБОУ ИВО «Московский государственный гуманитарно-
экономический университет»**

Направление «Профессиональное завтра в науке»

Номинация «Научная статья»

**«Система управления персональным компьютером, основанная на
использовании технологии eyetracking»**

Выполнил:
Самуйлов Константин
Александрович

Руководитель:
старший преподаватель
Савельева Оксана Николаевна

Москва, 2023

В последнее время растет интерес к развитию естественного взаимодействия между человеком и компьютером. Актуальность вопросов распознавания глазных жестов для помощи пользователям с ограниченными физическими возможностями определяется необходимостью создания условий и возможностей организации коммуникативного взаимодействия с внешней средой. Традиционно человек использует клавиатуру и мышь для управления интерфейсом компьютера, однако данные манипуляции сильно затрудняются или становятся вовсе невозможными, если у пользователя есть физические ограничения. Была поставлена задача по распознаванию жестов, произведенных глазами людей, страдающих поражением верхних конечностей. Данное программное обеспечение должно облегчить взаимодействие пользователя с компьютером.

Большинство современных пользовательских интерфейсов обеспечивают гораздо большую пропускную способность от компьютера к пользователю, чем в противоположном направлении. Графика, анимация, аудио и другие средства массовой информации могут быстро выводить большие объемы информации, но вряд ли существуют какие-либо средства ввода сравнительно больших объемов информации от пользователя. Таким образом, современные диалоги между пользователем и компьютером, как правило, довольно односторонни. Новые устройства ввода и носители информации, использующие «легкие» пассивные измерения, могут помочь устранить этот дисбаланс за счет удобного и быстрого получения данных от пользователя. Таким образом, движения глаз пользователя могут обеспечить удобный источник дополнительного пользовательского ввода с высокой пропускной способностью.

Разработка программного обеспечения для распознавания глазных жестов в помощь людям с ограниченными физическими возможностями необходима для создания средств, условий и возможностей организации для коммуникативного взаимодействия с внешней средой.

Использование компьютера в нашем мире доступно практически каждому. Однако, для людей с поражением верхних конечностей это становится серьезной проблемой. Такие операции как набор текста с клавиатуры, использование мыши, в зависимости от типа заболевания, становятся либо не возможными, либо значительно снижается скорость данных действий.

Одним из решений данной проблемы является создание программного обеспечения, позволяющего проводить манипуляции с персональным компьютером при помощи глазных жестов.

Ввод с помощью взгляда обладает рядом преимуществ, но также имеет и некоторые недостатки.

Преимущества:

- помогает людям с заболеваниями верхних конечностей при наборе текста, управлении компьютерной мышью;
- интуитивно понятен;
- каждому жесту можно присвоить любое действие;
- ускоряет взаимодействие пользователя с компьютером.

Недостатки:

- возможны ложные срабатывания жеста;
- конечность набора распознаваемых жестов.

Айтрекеры существуют уже несколько лет, но их использование в основном ограничивалось лабораторными экспериментами. Оборудование постепенно становится достаточно надежным и недорогим, чтобы рассматривать возможность использования в реальных интерфейсах пользователя и компьютера.

Самым простым решением было бы заменить мышь непосредственно на eyetracker — установить eyetracker и использовать его выходной поток x, y вместо потока мыши. Изменения в линии взгляда пользователя непосредственно приведут к перемещению курсора мыши. Но движение глаз отличается от движения рук.

Существуют значительные различия между ручным источником ввода, таким как мышь, и положением взгляда, которые необходимо учитывать при разработке методов взаимодействия, основанных на движении глаз:

- ввод движения глаз происходит значительно быстрее, чем в других современных средствах ввода. Прежде чем пользователь приведет в действие какое-либо механическое указывающее устройство, он обычно смотрит на пункт назначения, к которому хочет переместиться. Таким образом, движение глаз доступно в качестве указания цели пользователя до того, как он сможет активировать любое другое устройство ввода;

- «управление» глазом не требует никакой подготовки или особой координации для обычных пользователей; они просто смотрят на объект;

- глаз — это, конечно, гораздо больше, чем просто инструмент для высокоскоростного позиционирования курсора. В отличие от любого другого устройства ввода, eyetracker также показывает, на чем сосредоточен интерес пользователя. Самим действием наведения с помощью этого устройства пользователь меняет фокус своего внимания, и каждое изменение фокуса доступно компьютеру в виде команды наведения. Ввод с помощью мыши просто сообщает системе, что пользователь намеренно поднял мышь и направил ее на что-то. Ввод данных eyetracker может быть интерпретирован таким же образом (пользователь намеренно направил глаз на что-то, потому что он был обучен управлять этой системой таким образом). Но это также может быть истолковано как указание на то, на что пользователь в данный момент обращает внимание, без каких-либо явных действий ввода с его стороны;

- движение глазами часто является почти подсознательным действием. В отличие от мыши, относительно трудно постоянно сознательно и точно контролировать положение глаз. Взгляд постоянно перебегает с места на место, даже когда его владелец думает, что он пристально смотрит на один объект, и нежелательно, чтобы каждое такое движение инициировало компьютерную команду;

- в отличие от мыши, движение взгляда всегда «включено». Нет естественного способа указать, когда следует задействовать устройство ввода, как это происходит при захвате или отпуске мыши. Закрывание глаз отвергается по очевидным причинам - даже при отслеживании глаз в качестве входных данных основная функция глаз в диалоге между пользователем и компьютером заключается в общении с пользователем. Движения глаз являются примером более общей проблемы со многими новыми пассивными или некомандными средствами ввода, требующими либо тщательного проектирования интерфейса, чтобы избежать этой проблемы, либо какой-либо формы явного «сцепления» для включения и выключения мониторинга;

- оборудование для слежения за глазами по-прежнему гораздо менее стабильно и точно, чем большинство устройств ручного ввода.

Устройства слежения за глазами исторически делились на два типа. Первый является пассивным и ориентирован на определение взгляда пользователя относительно остального мира и, в частности, на какие элементы видимого поля в данный момент фокусируются. Второй более активен и рассматривает глаз не просто как средство наблюдения, но и как средство контроля.

В методе используются обычные камеры с видимым светом и методы компьютерного зрения для извлечения деталей о положении различных интересных объектов. Рост области компьютерного зрения за последние десять-пятнадцать лет привел к появлению множества методов, способных выполнять такой анализ.

Одним из преимуществ этого метода является то, что он не зависит от характеристик, которые чрезвычайно специфичны для глаза (например, градиенты заряда сетчатки или отражение инфракрасного зрачка), и может быть адаптирован к другим особенностям более сложных взаимодействий.

Система слежения является первым шагом в устройстве слежения за глазами и находится в непосредственном контакте с пользователем.

Отслеживание обрабатывает все данные с входных данных камеры, представляет пользователю интерфейсы калибровки и настройки параметров и выполняет алгоритмы, связанные с компьютерным зрением, для определения местоположения зрачков и бровей пользователя. С точки зрения пользовательского интерфейса, работа устройства выглядит следующим образом:

- Когда программа запускается, пользователю предлагается правильно расположить камеру, чтобы на видео были хорошо видны оба глаза.

- Затем пользователю будет предложено выбрать соответствующие области глаз из видео. Выбор области глаз очень важен, поскольку он позволяет нам значительно сузить пространство поиска, гарантировать, что на изображении отображаются только объекты, связанные с глазами, и сократить внешние помехи.

- После установки параметров начинается режим калибровки. На этом этапе пользователю предлагается выполнить несколько движений глазами и бровями во время тренировки программы. Конкретные требуемые действия — это смотреть влево, вправо, вверх и вниз и поднимать брови. После завершения калибровки пользователь готов приступить к использованию устройства. Устройство можно перекалибровать или отключить в любое время простым нажатием клавиши.

Данный алгоритм обработки глазных жестов имеет значительные минусы:

- низкая точность отслеживания зрачков, возможны ложные срабатывания;

- низкий контраст между чертами глаз/лица делает систему слежения неработоспособной. Другими словами, людей с карими/темными зрачками и/или бровями, близко подобранными к их тону кожи, трудно отследить из-за небольших различий в контрасте между этими чертами.

С целью решения поставленной задачи была разработана нейронная сеть, распознающая глазные жесты. Принцип её работы следующий:

- При помощи отдельной подпрограммы собираются данные о произведённых глазных жестах (рис.1);

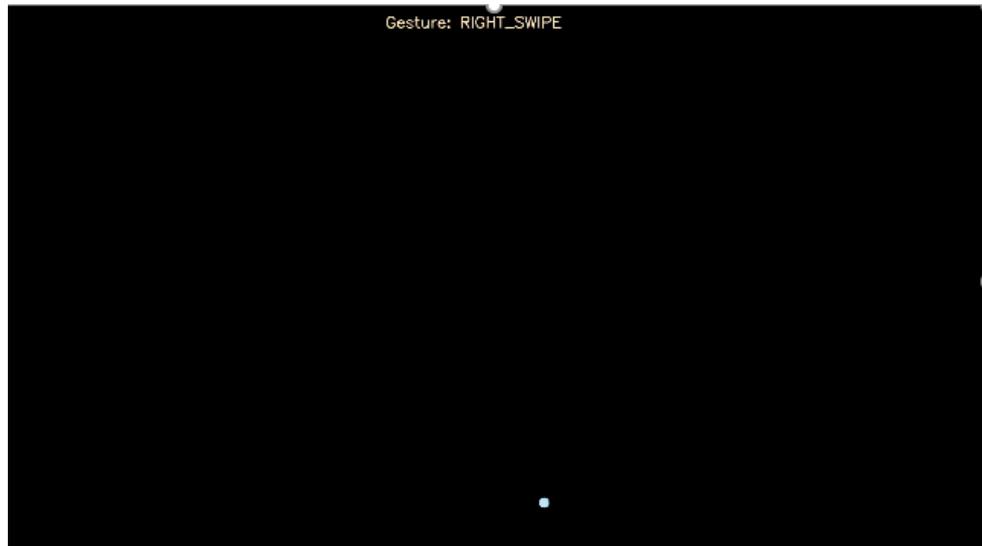


Рисунок 1. Сборщик данных

- Далее производится предобработка полученных данных, их структуризация и форматирование (рис.2);

	A	B	C	D	E	F	G
64	1650449661.91662	0.3833333333333333	0.0888888888888889	0.392555117607117	0.121794700622559	RANDOM_TELEPORT	1
65	1650449662.01663	0.3833333333333333	0.0888888888888889	0.392555117607117	0.121794700622559	RANDOM_TELEPORT	1
66	1650449662.11664	0.3833333333333333	0.0888888888888889	0.397460341453552	0.092642784118652	RANDOM_TELEPORT	1
67	1650449662.21665	0.3833333333333333	0.0888888888888889	0.380452632904053	0.086503505706787	RANDOM_TELEPORT	1
68	1650449662.31663	0.3833333333333333	0.0888888888888889	0.380452632904053	0.086503505706787	RANDOM_TELEPORT	1
69	1650449662.41666	0.3833333333333333	0.0888888888888889	0.36566948890686	0.129616022109985	RANDOM_TELEPORT	1
70	1650449662.51659	0.3833333333333333	0.0888888888888889	0.369199275970459	0.070502996444702	NONE	1
71	1650449662.61673	0.7166666666666667	0.7222222222222222	0.369199275970459	0.070502996444702	LEFT_SWIPE	1
72	1650449662.7166	0.7166666666666667	0.7222222222222222	0.375386595726013	0.049596548080444	LEFT_SWIPE	1
73	1650449662.81658	0.7166666666666667	0.7222222222222222	0.385972738265991	0.038660049438477	LEFT_SWIPE	1
74	1650449662.91691	0.7166666666666667	0.7222222222222222	0.389734029769898	0.029213786125183	LEFT_SWIPE	1
75	1650449663.0165	0.7166666666666667	0.7222222222222222	0.487257599830627	0.25463879108429	LEFT_SWIPE	1
76	1650449663.11654	0.7166666666666667	0.7222222222222222	0.628797769546509	0.629145860671997	LEFT_SWIPE	1
77	1650449663.2165	0.7166666666666667	0.7222222222222222	0.679898500442505	0.728026866912842	LEFT_SWIPE	1
78	1650449663.31656	0.7166666666666667	0.7222222222222222	0.72405195236206	0.75112247467041	LEFT_SWIPE	1
79	1650449663.41654	0.55	0.731481481481481	0.745896339416504	0.760881662368774	LEFT_SWIPE	1
80	1650449663.51657	0.3833333333333333	0.740740740740741	0.753878593444824	0.73829710483551	LEFT_SWIPE	1
81	1650449663.61652	0.2166666666666667	0.75	0.753878593444824	0.73829710483551	LEFT_SWIPE	1
82	1650449663.71652	0.05	0.759259259259259	0.755066514015198	0.727370500564575	LEFT_SWIPE	1
83	1650449663.81661	-0.1166666666666667	0.768518518518519	0.518730044364929	0.70750904083252	LEFT_SWIPE	1
84	1650449663.91668	-0.2833333333333333	0.777777777777778	0.518730044364929	0.70750904083252	LEFT_SWIPE	1
85	1650449664.0166	-0.45	0.787037037037037	0.308265447616577	0.730538845062256	LEFT_SWIPE	1
86	1650449664.11661	-0.6166666666666667	0.796296296296296	-0.101232886314392	0.751424193382263	LEFT_SWIPE	1
87	1650449664.21668	-0.7833333333333333	0.805555555555556	-0.101232886314392	0.751424193382263	LEFT_SWIPE	1
88	1650449664.31668	-0.95	0.814814814814815	-0.3705052113737488	0.776047945022583	LEFT_SWIPE	1
89	1650449664.41662	-1.1145833333333333	0.824074074074074	-0.474201321601868	0.774099707603455	NONE	1
90	1650449664.51697	0.8197916666666667	0.774074074074074	-0.590232014656067	0.772661566734314	LEFT_SWIPE	1
91	1650449664.61658	0.8197916666666667	0.774074074074074	-0.626240104436874	0.779057336807251	LEFT_SWIPE	1
92	1650449664.71657	0.8197916666666667	0.774074074074074	-0.856130853295326	0.803384184837341	LEFT_SWIPE	1
93	1650449664.81658	0.8197916666666667	0.774074074074074	-0.929284282028675	0.801417827606201	LEFT_SWIPE	1
94	1650449664.91658	0.8197916666666667	0.774074074074074	0.24115264415741	0.948173522949219	LEFT_SWIPE	1
95	1650449665.01658	0.8197916666666667	0.774074074074074	0.623085975646973	0.913232564926147	LEFT_SWIPE	1
96	1650449665.11655	0.8197916666666667	0.774074074074074	0.819867372512817	0.805844068527222	LEFT_SWIPE	1
97	1650449665.21648	0.8197916666666667	0.774074074074074	0.819867372512817	0.805844068527222	LEFT_SWIPE	1
98	1650449665.31657	0.653125	0.762962962962963	0.831560611724854	0.799769401550293	LEFT_SWIPE	1
99	1650449665.41662	0.4864583333333334	0.751851851851852	0.831560611724854	0.799769401550293	LEFT_SWIPE	1
100	1650449665.5168	0.3197916666666667	0.742592592592593	0.831560611724854	0.799769401550293	LEFT_SWIPE	1
101	1650449665.61669	0.153125	0.731481481481481	0.868713617324829	1.27227592468262	LEFT_SWIPE	1
102	1650449665.71673	-0.0135416666666667	0.7222222222222222	0.780699133872986	0.989984154701233	LEFT_SWIPE	1
103	1650449665.81673	-0.1802083333333333	0.7111111111111111	0.498648405075073	0.881306171417236	LEFT_SWIPE	1
104	1650449665.91672	-0.346875	0.7	0.127965092658997	0.801228165626526	LEFT_SWIPE	1
105	1650449666.01673	-0.5125	0.690740740740741	-0.040708899497986	0.779116272926331	LEFT_SWIPE	1
106	1650449666.11669	-0.6791666666666667	0.67962962962963	-0.177733719348908	0.759712934494019	NONE	1
107	1650449666.2168	-0.5125	0.896296296296296	-0.33717942237854	0.743506908416748	RIGHT_SWIPE	1
108	1650449666.31652	-0.5125	0.896296296296296	-0.536303331158638	0.731434345245361	RIGHT_SWIPE	1
109	1650449666.41652	-0.5125	0.896296296296296	-0.670091688632965	0.728979229927063	RIGHT_SWIPE	1

Рисунок 2. Собранные данные

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Neural Network Tutorial – Multi Layer Perceptron <https://www.edureka.co/blog/neural-network-tutorial/> [Электронный ресурс] // URL: www.edureka.co/blog/neural-network-tutorial/
2. Understanding how Long Short-Term Memory works in classification of sequences of symbols. [Электронный ресурс] // URL: https://translated.turbopages.org/proxy_u/en-ru.ru.8dea32a0-62bad7ea-c865ad24-74722d776562/https://datascience.stackexchange.com/questions/107897/understanding-how-long-short-term-memory-works-in-classification-of-sequences-of
3. Understanding of LSTM Networks // GeeksforGeeks URL: <https://www.geeksforgeeks.org/understanding-of-lstm-networks/> (дата обращения: 13.02.2022).
4. Даг Хеллман Стандартная библиотека Python 3. Справочник с примерами. - 2-е изд. - М.,: Вильямс, 2019. - 1376 с.
5. Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018. – 121 с.
6. Как действует eyetracking? [Электронный ресурс] // URL: <https://www.eyetracking.care/informatsiya/>
7. Коберн Алистер, Современные методы описания функциональных требований к системам . - М.,: Лори, 2022. - 283 с.
8. Франсуа Шолле, Глубокое обучение на Python. — СПб.: Питер, 2018. — 400 с.
9. Юси Лю, Обучение с подкреплением на PyTorch: сборник рецептов / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2020. – 282 с.

зрачками и/или бровями, близко подобранными к их тону кожи, трудно отследить из-за небольших различий в контрасте между этими чертами.

После исследования был сделан вывод о том, что данный метод не является эффективным при управлении интерфейсом.

Описаны существующие варианты решения поставленной задачи. На основе теоретических знаний, разработано приложение, распознающее глазные жесты с помощью технологии eyetracking.

Разработано программное обеспечение для сбора и визуализации данных получаемых с eyetracker.

Результаты, полученные в ходе тестирования на реальных данных с eyetracker, подтвердили адекватность разработанной модели и практическую применимость для реализации зрительного интерфейса управления роботизированной инвалидной коляской.

- После структуризации массив данных передаётся в рекуррентную нейронную сеть (процесс обучения сети показан на рис.3);

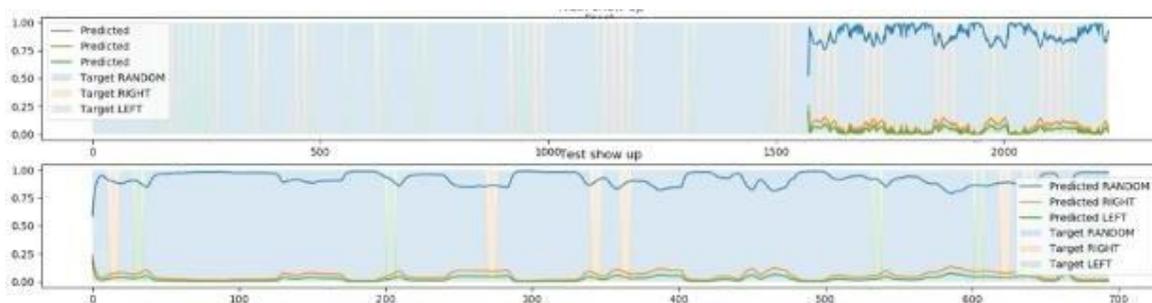


Рисунок 3. Обучение нейронной сети

- В итоге, при помощи обученной модели, мы можем определять глазные жесты, произведённые пользователем.

На рисунке 3 представлено обучение нейронной сети. Оранжевая заливка обозначает правый свайп, которому сеть должна обучиться. Зеленая заливка – левый свайп. Снизу представлен тестовый dataset, который показывает, насколько правильно обучается сеть. Синяя линия - произвольное движение точки. О правильно обученном жесте сигнализирует пересечение синей линии с линией жеста (зеленая и желтая).

ЗАКЛЮЧЕНИЕ

В ходе исследования был проведен анализ существующих способов и моделей, применяемых для глазоуправления, а также приборов, используемых в практической реализации зрительного интерфейса для управления роботизированной инвалидной коляской лицами с нарушением опорно-двигательного аппарата. Были рассмотрены методы, использующие обычные камеры с видимым светом и метод, основанный на технологии eyetracking. При исследовании технологии с использованием обычных камер выявлены следующие недостатки:

- низкая точность отслеживание зрачков, возможны ложные срабатывания;
- низкий контраст между чертами глаз/лица делает систему слежения неработоспособной. Другими словами, людей с карими/темными