

САНКТ-ПЕТЕРБУРГСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
Математико-механический факультет

Кафедра системного программирования

ЭОГ-метод управления  
компьютером при помощи глаз:  
обнаружение предсаккадных пиков

Дипломная работа студентки 545 группы

**Храмцовой Елены Александровны**

Научный руководитель	.....	разработчик ПО «Ланит-Терком»
	/подпись/	Пименов А. А.
Рецензент	.....	к.ф.-м.н., доц.
	/подпись/	Кознов Д. В.
«Допустить к защите»	.....	д.ф.-м.н., проф.
заведующий кафедрой	/подпись/	Терехов А. Н.

Санкт-Петербург

2010

SAINT PETERSBURG STATE UNIVERSITY

Mathematics & Mechanics Faculty

Software Engineering Chair

EOG eye-computer interface:  
presaccadic peak detection

by

**Elena Khramtsova**

Master's thesis

Supervisor	.....	software developer "Lanit-Tercom"
	/signature/	A. A. Pimenov
Reviewer	.....	Ass. Prof., PhD
	/signature/	D. V. Koznov
"Approved by"	.....	Prof., PhD
Head of Department	/signature/	A. N. Terekhov

Saint Petersburg

2010

# Содержание

<b>1</b>	<b>Введение</b>	<b>3</b>
<b>2</b>	<b>Обзор литературы</b>	<b>5</b>
2.1	Применение систем слежения за взглядом . . . . .	5
2.1.1	Неврология . . . . .	5
2.1.2	Психология . . . . .	6
2.1.3	Организация производства и человеческий фактор .	7
2.1.4	Маркетинг, рекламный бизнес и исследования usability . . . . .	8
2.1.5	Информатика и задача управления устройствами . .	8
2.2	Основные принципы устройства систем слежения за взгля- дом . . . . .	9
2.2.1	Анализ магнитного поля . . . . .	10
2.2.2	Слежение за взглядом с помощью видеокамеры . . .	11
2.2.3	Сравнение метода измерения магнитного поля и метода видеосъемки . . . . .	18
2.2.4	Слежение за взглядом с помощью ЭОГ . . . . .	18
2.2.5	Сравнение видеосистем и систем, основанных на ЭОГ-методе . . . . .	19
2.3	Интерфейсы мозг-компьютер . . . . .	21
2.4	Метод наименьших квадратов . . . . .	25
<b>3</b>	<b>Система «Окулографический интерфейс»</b>	<b>27</b>
3.1	Предназначение и функциональность системы . . . . .	27
3.2	Архитектура системы . . . . .	28
3.3	Проблемы . . . . .	31
3.4	Пути решения проблем . . . . .	32

<b>4</b>	<b>Техническая постановка задачи</b>	<b>35</b>
<b>5</b>	<b>Предлагаемые алгоритмы</b>	<b>36</b>
5.1	«Стихийный» алгоритм . . . . .	36
5.2	Адаптированный метод наименьших квадратов . . . . .	37
<b>6</b>	<b>Реализация</b>	<b>41</b>
6.1	Функциональность и интеграция в систему . . . . .	41
6.2	Особенности реализации . . . . .	42
<b>7</b>	<b>Тестовое окружение, тестирование и апробация</b>	<b>44</b>
7.1	Архитектура и средства реализации . . . . .	44
7.2	Методика тестирования . . . . .	45
<b>8</b>	<b>Результаты апробации</b>	<b>49</b>
<b>9</b>	<b>Заключение</b>	<b>54</b>

# 1 Введение

Анализ движений глаз на данный момент является широко распространенной задачей и имеет большое количество приложений в различных областях, таких как медицинская диагностика, маркетинговые исследования, исследования usability. В частности интересна задача управления устройствами при помощи глаз - это очень полезно людям с ограниченными физическими возможностями, а также тем, кто проводит много времени, например, за компьютером и устает действовать руками.

За последние 20 лет появилось много систем, распознающих движения глаз человека. Большинство существующих систем основаны на обработке видеозображения зрачка человека. Этот подход имеет как преимущества (при желании можно добиться высокой точности), так и недостатки (относительно высокая вычислительная сложность).

Менее распространенный подход используется в системе «Окулографический интерфейс», разрабатываемой на базе исследований механизма зрения в НИИ физиологии им. А. Ухтомского при биологическом факультете СПбГУ. «Окулографический интерфейс» решает задачу управления персональным компьютером при помощи движений глазами. Система является альтернативой компьютерной «мыши». Она основана на следующем подходе. На голове испытуемого размещают специальные электроды, которые соединяются с программно-аппаратной системой, позволяющей фиксировать изменения электрического потенциала при движении глаз - известно, что такие изменения достаточно заметны при движении глаз. Данные изменения анализируются, и в результате определяется направление движения глаз. Полученная информация используется, в частности, для управления курсором компьютер-

ной мыши глазами, что крайне востребовано для людей с ограниченными физическими возможностями. В данный момент система используется коллективом НИИ для исследований физиологии глазодвигательного аппарата в зависимости от пола, возраста и других особенностей человека [6]. Функциональность системы в этом случае сводится к визуализации электроокулограммы (ЭОГ). Кроме того, проводятся межфакультетские практикумы Психологического и Биолого-Почвенного факультета с целью изучения физиологии движений глаз и психологии процесса обучения.

В рамках физиологических исследований обнаружено, что перед каждым движением зрачка происходит быстрое напряжение окологлазных мышц, что вызывает короткий резкий скачок потенциала на всех каналах электроокулограммы - предсаккадный пик. Это означает, что существует возможность сделать вывод о движении взора еще до начала самого движения.

Целью данной работы является разработка и реализация метода обнаружения предсаккадных пиков с учетом их одновременного (или с малой задержкой) появления на разных каналах, реализация соответствующей компоненты и интеграция в систему «Окулографический интерфейс». Кроме того, поставлена задача реализации блока сохранения амплитуды саккады в рамках разработки дискретизирующего фильтра, позволяющего уменьшить время реакции системы на действия пользователя.

## 2 Обзор литературы

### 2.1 Применение систем слежения за взглядом

Системы отслеживания взора предназначены для определения позиции глаз и нахождения параметров их движения. Такие системы применяются во многих областях [11]. Расскажем о них подробнее.

#### 2.1.1 Неврология

В неврологии анализ движений глаз широко применяется как экспериментальный метод исследования механизма зрения. Так, благодаря этим работам, известно, какие именно области мозга задействованы в процессе зрения, какова психологическая организация, когнитивные и поведенческие аспекты зрения. На данный момент исследования ведутся в следующих направлениях:

- *Неврология внимания.* Например, изучается активность нейронов обезьяны с целью ответа на вопрос, какие именно нейроны задействованы во время фиксации взгляда [19].
- *Движения глаз и генерация изображения в мозгу.* В таких исследованиях одновременно следят за движением глаз и активностью мозга во время выполнения разных заданий [15].
- *Диагностика.* С помощью слежения за взглядом можно диагностировать неврологические расстройства или усталость (например, для спортсменов).

### 2.1.2 Психология

В области психологии изучается визуальное внимание во время выполнения заданий по обработке информации. Опишем основные направления исследований.

- *Процесс чтения.* Во-первых, активно исследуется направление, скорость и амплитуда движений глаз при чтении, для чего достаточно опосредованной обработки записей. Так, известно, что при чтении английского текста фиксации взора длятся 200-250 мс, а размер саккад<sup>1</sup> в 7-9 раз превышает размер буквы. Все это зависит от содержания текста и его типографских параметров.
- *Восприятие сцены.* В отличие от чтения, нет канонического направления сканирования (например, слева направо, или сверху вниз). Особый интерес представляет восприятие картин, фильмов.
- *Зрительный поиск.* Исследователи пытаются ответить на вопрос о том, как именно двигаются глаза в процессе поиска в тексте, на картинке, в массивах случайных букв и цифр.
- *Естественные задания.* Интересно также, как движутся глаза в обычных ситуациях, например, сделан такой вывод: несмотря на то, что некоторые действия, например, приготовления чая, являются «автоматическими» и почти не задействуют сознание, глаза тщательно контролируют каждую стадию процесса [14]. Для таких исследований необходимы системы слежения за взглядом, не ограничивающие движения глаз, головы и рук.

---

<sup>1</sup>Саккады (от франц. «хлопок паруса») – скачкообразные быстрые содружественные фиксирующие движения глаз. Возникают, когда взгляд переводится с одного неподвижного предмета на другой, перемещаясь одним движением [1].



- *Обработка устной речи.* Здесь исследуются движения глаз в течении одновременного прослушивания истории и просмотра некоторой информации.

### 2.1.3 Организация производства и человеческий фактор

Очень важно иметь возможность оценить условия, в которых люди работают или будут работать. Традиционные методы включают в себя измерение времени и точности реакции, то есть как быстро человек выполняет задачу, и насколько качественно. Чтобы изучить отдельные шаги, нужно анализировать выполнение каждой процедуры. Для этого также весьма полезен анализ движений глаз, так как он дает дополнительные сведения о зрительных, когнитивных и относящихся к вниманию аспектах выполнения задач. Рассмотрим области, где это особенно интересно и распространено.

- *Авиация.* Технологии отслеживания взгляда используются в симуляторах полета [7]. При этом оценивается удобство (usability) некоторых новых инструментов, например, насколько удобно пользоваться новыми электронными картами, и т.д. Также об уже существующем оборудовании, полезно знать, куда чаще всего смотрит пилот, откуда ему удобнее брать информацию. Исследования показали, что анализ движений глаз подходит для оценки профессионализма пилота и тренировки будущих пилотов.
- *Возждение автомобиля.* Известно, что большое число дорожно-транспортных происшествий происходит из-за дефицита зрительного внимания (сонливость за рулем). При таком состоянии можно наблюдать длительную фиксацию взгляда в одной точке. Соответ-

ственно, методом анализа движений глаз легко находить такие состояния и будить водителя, например, громкими звуками. Так же, как и в авиации, здесь можно использовать технологии слежения за взглядом в различных симуляторах.

#### **2.1.4 Маркетинг, рекламный бизнес и исследования usability**

Слежение за взглядом может помочь в таких областях, как тестирование эффективности рекламы (copy testing) в печати, изображениях, видео или графике и т.д. Можно получить информацию о том, как покупатель распределяет зрительное внимание между разными формами рекламы. Кроме того, движения глаз проливают свет на то, как именно человек использует графический интерфейс. Так можно получить информацию о том, как организовать эффективные выпадающие меню или web-страницу. Крупные компании заинтересованы в таких исследованиях своих программных продуктов, и сами или через партнеров такие исследования проводят. Так, например, в компании Google есть группа, занимающаяся исследованиями в данной области (Google's eye-tracking team [16]).

#### **2.1.5 Информатика и задача управления устройствами**

Систему слежения за взглядом можно использовать в качестве устройства ввода, например, как аналог компьютерной мыши. Такое устройство ввода может понадобиться людям с ограниченными физическими возможностями, а также тем, кто проводит очень много времени за компьютером и устает действовать руками, например, людям, регулярно играющим в компьютерные игры. Также известно, что если объекты в среде сильно удалены друг от друга, то управление глазами

намного быстрее, чем управление руками [20].

Первоначальные приложения, основанные на методе анализа движений глаз - набор текста глазами. Такие системы были и являются очень полезными для некоторых людей с ограниченными возможностями. Обычно такие системы показывают пользователю виртуальную клавиатуру на мониторе или спроецированную на стену. По направлению взгляда система узнает, какая именно буква имелась в виду, а по задержке - выбрана она, или нет. Система дает ответ пользователю визуально, звуками, или и тем, и тем. Проблемы здесь следующие: если устройство недостаточно точное, а особенно при ошибке в калибровке, оно будет выдавать неверный ответ, также не понятно, как обеспечить аналог нажатия на кнопку мыши (Midas Touch problem).

Существуют системы, в которых «глазное» управление корректируется «ручным».

## **2.2 Основные принципы устройства систем слежения за взглядом**

По принципу обработки системы делятся на три группы: анализ видеоизображения зрачка, исследование магнитного поля в окрестности глаза, и анализ электроокулограммы.

Системы слежения за взглядом могут либо собирать информацию и потом анализировать записи (offline обработка), либо анализировать данные по мере их поступления (online системы). Расскажем о возможностях применения этих классов систем на примере исследований в психолингвистике [3].

В психолингвистике используется два типа экспериментальных методов: опосредованные (исследуют процесс понимания после того, как

он уже завершился) и непосредственные (исследуют процесс понимания по мере того, как он происходит). Преимущества непосредственных методов очевидны: возможна запись времени реакции испытуемого, что позволяет измерить нагрузку на аппарат понимания. С помощью устройств, основанных на таких методах, регистрируются и измеряются следующие характеристики когнитивного процесса, вызывающего движения глаз: длина саккад (обычно это 30 – 50 мс), и длина периодов фиксации (200 – 300 мс).

### 2.2.1 Анализ магнитного поля

Согласно [13], системы, основанные на анализе магнитного поля, устроены следующим образом. В глаз вставляется кремниевое кольцо, содержащее стержень из тонкого медной проволоки. Если поместить испытуемого в магнитное поле переменного тока, положение глаза может быть вычислено по уровню напряжения, возникающего внутри кольца. Этот метод дает очень точные результаты, уровень шума низок, это позволяет производить измерения с большим пространственным ( $\ll 1^\circ$ ) и временным ( $\ll 1$  мс) разрешением, поэтому в 1998 году данный метод признан стандартом *de facto* в окуломоторных исследованиях. Также он используется в исследованиях психологии движений глаз.

Слабой стороной этого подхода является необходимость неприятных инвазивных процедур. Даже с анестезией, большинство людей не выдерживает надетое на глаз проводящее кольцо больше 30 минут. Человек едва ли будет согласен использовать такие методы, если только это не жизненная необходимость, или ему не платят заработную плату как участнику эксперимента. Кроме того, применение подобных методов невозможно по отношению к детям и больным.

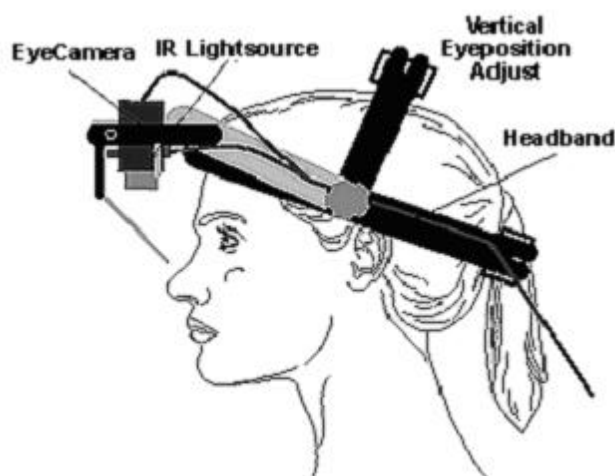


Рис. 1: Конструкция видеосистемы [12]

### 2.2.2 Слежение за взглядом с помощью видеокамеры

Этот тип систем слежения за взором основан на использовании бесконтактных оптических методов. Свет, обычно инфракрасного диапазона, отражается от глазного яблока и улавливается видеокамерой или другим специальным оптическим сенсором. Далее, из анализа изменений отражения делается вывод о повороте глаза (см. рис. 1).

Такие системы следят за отражением от внешней поверхности роговицы (первым изображением Пуркинье) и центром зрачка во времени. Более чувствительные системы (dual-Purkinje eye trackers) отслеживают первое изображение Пуркинье и отражение с внутренней стороны хрусталика (четвертое изображение Пуркинье).

По данным [2], регистрация окуломоторной активности может выполняться как бинокулярно, так и монокулярно. Бинокулярные установки бывают «истинными» (предоставляющие данные о положении каждого глаза в отдельности). Они используются, например, при изучении

вергентных движений глаз, сопряженности движений глаз при чтении и восприятия глубины. Поскольку для большинства людей характерно синхронное движение обоих глаз, для многих исследовательских и прикладных задач подходят монокулярные установки, которые являются более простыми, а значит более дешевыми. Другой класс установок можно назвать «псевдобинокулярными». Они проводят видеосъемку обоих глаз, но выдают усредненные данные.

Важными характеристиками являются точность регистрации и разрешающая способность. *Точность регистрации* - мера соответствия воспринимаемого объекта и линии зрения - теоретически она ограничена размерами фовеа<sup>2</sup>, т.е.  $1,5 - 2^\circ$ . Однако фактически, если в процессе калибровки каждая из калибровочных меток будет иметь одну и ту же проекцию на фовеа, точность регистрации может достигать менее  $0,5^\circ$ . На значение точности влияет ряд факторов:

1. различие в окуломоторном поведении испытуемого в процессе калибровки и во время самостоятельной работы
2. смещение системы координат зрения в ходе исследования
3. факторы, в целом влияющие на регистрацию движений глаз (ношение очков или линз испытуемым, индивидуальные особенности физиологии глаза испытуемых, уровень освещенности и т.д.)

*Разрешающая способность регистрирующих систем* - насколько мелкие движения глаз могут детектироваться. Это количество пикселей получаемого изображения, приходящихся на область зрачка. Лучшие системы

---

<sup>2</sup>В задней части глаза, где зрительная ось пересекает сетчатку, имеется углубление - фовеа, обильно населенная колбочками, отвечающими за зрение при дневном освещении [1]

имеют разрешение порядка  $0,01^\circ$ , что позволяет регистрировать микро-саккады и дрейф.

Для систем, основанных на видеосъемке, также крайне важна *частота дискретизации* видеокамеры (sampling rate) или, другими словами, та скорость, с которой происходит съемка. Системы можно разделить на:

1. очень медленные - частота дискретизации 25 Гц,
2. среднескоростные - 50-100 Гц,
3. скоростные - 250-500 Гц;
4. высокоскоростные видеосистемы регистрации - 1000-2000 Гц

В последнем случае можно получить данные о времени пребывания взора в зоне интереса, при этом продолжительность фиксации будет иметь относительную точность с колебаниями  $\pm 20$  мс. Точность определения продолжительности фиксаций для среднескоростных систем составляет  $\pm 10$  мс, эти системы могут вполне эффективно использоваться в большинстве прикладных исследований. Однако более «тонкие» характеристики движений глаз, например, такие как скорость саккады, не могут быть выявлены с помощью данного типа систем регистрации. Скоростные системы имеют возможность детектировать практически большинство «нюансов» окуломоторной активности человека. Как правило, они предполагают устойчивую фиксацию головы испытуемого. Высокоскоростные системы помимо еще большей точности и разрешающей способности предоставляют исследователю возможность проводить on-line анализ окуломоторной активности человека и регулировать параметры предъявляемых стимулов.

Видеокамеры вообще довольно дешевы, но при ужесточении требований к ним - при увеличении частоты дискретизации - качественно меняется принцип их работы, а значит существенно возрастает стоимость.

По данным компании Arrington Research [8], системы с жесткой фиксацией головы стоят от \$6500 (монокулярные системы) и от \$8700 (бинокулярные). Системы, в которых камера крепится к голове пользователя, стоят от \$9500 монокулярные, и от \$12000 бинокулярные. Верхний же предел цены для таких устройств составляет \$71000.

Системы отслеживания взгляда, основанные на видеосъемке, пользуются большой популярностью, так как являются неинвазивными - не требуют нарушения естественных внешних барьеров организма. А значит, при их использовании нет риска травмы или заражения, который сопутствует любому более менее серьезному врачебному вмешательству. Такими устройствами человек может пользоваться самостоятельно в повседневной жизни, а не только в лаборатории в рамках некоторых исследований.

Популярность видеосистем, спрос на них влечет за собой обилие коммерческих реализаций представленных выше методик. Недостатком этого класса систем отслеживания движений глаз является их высокая стоимость, которая автоматически лишает некоторую часть потенциальных потребителей возможности их использовать.

Приведем сравнительное описание нескольких продуктов фирмы SensoMotoric Instruments (SMI) [2] семейства *SMI iView X*.



Установка	Hi-Speed	RED	HED
Монокулярная	+	-	+
Биноккулярная	+	+	-
Фиксация головы	+	+-	-
Разрешение	$\leq 0,01^\circ$	$0,1^\circ$	$0,1^\circ$
Дискретизация (Гц)	1250 или 2.500	50	50 или 200
Точность	$0,25^\circ - 0,5^\circ$	$0,5^\circ \sim 1,0^\circ$	$0,5^\circ - 1,0^\circ$
Примечание			

*Установка SMI iView X Hi-Speed - 1250/500* состоит из компьютера с двумя мониторами и регистрирующей стойки-колонны. *Установка iView X RED* предоставляет возможность движения головы и верхнего плечевого пояса испытуемого (правда, весьма ограниченную). Данная установка располагается на столе и состоит из компьютера с двумя мониторами и небольшого портативного блока, который расположен под экраном монитора. Второй монитор используется исследователем для целей контроля. *Установка iView X HED* - крепится на велосипедном шлеме или на бейсболке с жестким козырьком и позволяет осуществлять видеорегистрацию окуломоторной активности при свободном поведении испытуемого (например, покупка какого-то товара в магазине, вождение автомобиля и т.д.). Имеется ряд моментов, ограничивающих применение установки. Во-первых, точность производимой регистрации. Во-вторых, метрика пространства, в котором будет осуществляться свободное поведение испытуемого, в значительной мере может отличаться от метрики, в которой была произведена калибровка. Кроме того, поскольку излучатель и камера находятся вблизи от глаза, они существенно ограничивают обзор. Это толкает испытуемых на неестественное вскидывание головы. Также часто при экстремальных амплитудах глаз уходит из зоны видео-

регистрации или зоны инфракрасной подсветки, и данные о направлении взгляда теряются.

Существует дополнение этой системы излучателем который образует электромагнитное поле радиусом 1,5-2 м. Установка *iView X HED+HT* состоит из мобильной системы регистрации, которая монтируется на велосипедном шлеме с датчиком, чувствительным к изменению электромагнитного поля для отслеживания движений головы. До начала исследования экспериментатор должен точно задать базисные прямоугольные области в пространстве электромагнитного поля. Основным недостатком данной системы - затраты большого количества времени на наладку оборудования, высокая чувствительность к внешним электромагнитным влияниям и к перемещению излучателя.

### **Классические алгоритмы выделения фиксаций и саккад**

- Алгоритм детекции саккад, основанный на определении угловой скорости [18].

Система интерпретирует движение как саккаду в том случае, если скорость движения в середине пути от одной точки к другой превышает определенное пороговое значение, например,  $75^\circ/\text{сек}$ . В противном случае система детектирует фиксацию. При достаточной частоте дискретизации (более 200 Гц) этот алгоритм дает возможность распознать короткие саккады, которые очень часто встречаются, например, при чтении.

Алгоритм, основанный на определении угловой скорости, чувствителен к величине нижнего порога угловой скорости. С его увеличением растет средняя продолжительность обнаруживаемых фиксаций (так, при значении порога в  $90^\circ/\text{сек}$  средняя продолжитель-

ность фиксации возрастает в 2,5 раза по сравнению со значением порога в  $30^\circ/\text{сек}$ ), уменьшается их количество, а также снижается общее количество детектируемых саккад.

- Алгоритм детекции фиксации, базирующийся на определении дисперсии и продолжительности фиксации [17].

Фиксацией считается набор соседних отсчетов, находящихся внутри пространства заданного диаметра, общей продолжительностью не меньше заданной. Вариация управляющих параметров существенно влияет на результат. Так, средняя продолжительность фиксации при временном пороге в 100 мс и при пороге дисперсии в 100 пикселей в 2,5–3 раза выше, чем при том же самом временном пороге, но при дисперсии в 20 пикселей.

- Алгоритм детекции, основанный на определении угловой скорости, использующий дополнительный критерий продолжительности фиксации.

Данный алгоритм использует принципы детекции движений глаз из двух вышеуказанных алгоритмов. Правда, во многом он представляет собой лишь гипотетическую модель, не апробированную в ходе исследований.

Для высокоскоростных систем регистрации с высоким разрешением более продуктивно обращаться к алгоритму, основанному на определении угловой скорости. Для систем с небольшой частотой дискретизации подходит алгоритм детекции фиксации. Все рассмотренные алгоритмы хорошо работают в случае фиксированного положения головы испытуемого относительно предъявляемого изображения и применимы преимущественно для случаев рассматривания статических изображений.

### **2.2.3 Сравнение метода измерения магнитного поля и метода видеосъемки**

В статье [13] сравнивается подход к проблеме распознавания движений глаз, основанный на видеонаблюдении, с исследованием магнитного поля. Вывод заключается в том, что системы с проводящим кольцом существенно точнее, чем видеосистемы, хотя есть вероятность, что при использовании первых глаза движутся не как обычно из-за дискомфорта.

### **2.2.4 Слежение за взглядом с помощью ЭОГ**

Последний тип систем слежения за взглядом - тот, к которому принадлежит и разрабатываемая система - основан на обработке электроокулограммы (в дальнейшем ЭОГ). Расскажем о нем подробнее.

Электроокулография является отдельной отраслью электрофизиологии. Согласно [9], глазное яблоко представляет собой электрический диполь, где сетчатка является отрицательным полюсом, а роговица - положительным. Значит, глаза являются источниками потенциального электрического поля, которое можно уловить даже в абсолютной темноте или при закрытых веках. Когда глаз движется от центра к периферии, сетчатка приближается к одному электроду, а роговица - к другому. Это изменение ориентации диполя, а значит и в электрическом потенциальном поле, дает в результате изменение измеряемого сигнала. Это происходит при скачкообразных смещениях взора (или саккадах), при глазном треморе для дезадаптации фоторецепторов (или микросаккадах), при медленном смещении взора при визуальном слежении или наклонах головы, а также при морганиях. Знак отклонения потенциала зависит от расположения электрода и от направления поворота глаз, а амплитуда

- от угла поворота глазного яблока. Путем анализа этих изменений сигнала можно восстановить направление и амплитуду движения глаза. По всей видимости, сейчас это является довольно грубым методом измерения саккадических движений, соответствующих резким сдвигам взгляда и обнаружения морганий.

Систем отслеживания взгляда, основанных на ЭОГ-методе, в данный момент очень мало. Наиболее известной является *Wearable EOG Goggles*. Эта система разработана группой Андреаса Буллинга (Andreas Bulling) в апреле 2009 года [10]. Выделим основные принципы детекции саккад.

- Для детекции саккад используется разложение сигнала по вейвлетам Хаара.
- Однонаправленные саккады, идущие друг за другом с интервалом не более 0.06 секунды (число взято из биологических исследований), интерпретируются как одну саккаду.
- Направление саккад преобразуются в символы (по принадлежности к 16 секторам на окружности), последовательность саккад таким образом переходит в строку, в которой далее ищутся определенные шаблоны, соответствующие глазным «жестам».

### **2.2.5 Сравнение видеосистем и систем, основанных на ЭОГ-методе**

Анализ ЭОГ - это облегченный подход, который, в противоположность видеосистемам, требует очень мало вычислительной мощности, работает в любых световых условиях. Он может быть реализован в качестве отдельного переносного устройства, которое может применяться для детекции движений глаз в повседневной жизни, или для распознавания

фазы быстрого сна, которая характеризуется повышенной активностью головного мозга, а ее признаком являются интенсивные движения глазных яблок с возрастающей амплитудой.

Также преимущества систем, основанных на регистрации ЭОГ, перед видеосистемами состоит в следующем. По данным [2], видеооборудование оптимально работает на испытуемых с нормальным зрением, не корректируемым с помощью очков или контактных линз и не использующих тушь для ресниц. Очки (в особенности бифокальные), и в меньшей мере контактные линзы, вызывают появление дополнительных бликов, мешающих определению направления взгляда. Накрашенные ресницы, увеличиваясь в объеме, частично перекрывают поле съемки видеокамеры, особенно когда взгляд испытуемого направлен в нижнюю часть экрана; помимо этого, тушь может отражать инфракрасный свет и, таким образом, дезориентировать регистрирующую систему новым бликом. Возможное решение указанных проблем состоит в выборе оптимального угла наклона регистрирующей камеры и исключении из анализа части снимаемого камерой кадра. Далее, в таких экстремальных условиях, как задымление (например, при пожарах), повышенная или переменная влажность (например, в горах), оптика будет плохо работать. Кроме того, устройства с инфрокрасной подсветкой и камерой плохо работают со светлыми (голубыми) глазами, при ярком дополнительном освещении или на солнце.

Системы же, основанные на ЭОГ, абсолютно кго всему этому нечувствительны, важно только, чтобы пользователь имел способность двигать глазами. Даже при потере зрения во взрослом возрасте, организм часто сохраняет эту способность.

Основное различие этих двух типов систем заключается в том, что видеосистемы анализируют результат когнитивных процессов, приводя-

щих к движениям глаз, смотрят как бы «извне». А те, что основаны на ЭОГ, способны в некоторой мере анализировать механизмы, приводящие к движению глаз как к результату - так, известно, что намерение совершить движение глазами отражается на ЭОГ в виде небольшого пика, предшествующего саккаде, причем не зависимо от направления и амплитуды скачка, этот пик всегда идет на всех каналах и направлен одинаково. Поэтому существует возможность судить о наличии движения еще до его начала.

Систем, основанных на ЭОГ, сейчас очень мало.

### **2.3 Интерфейсы мозг-компьютер**

Система «Окулографический интерфейс» предназначена для управления персональным компьютером, то есть является интерфейсом глаз-компьютер. Родственной задачей является построение так называемого интерфейса мозг-компьютер(ИМК).

Существует довольно много методов решения данной задачи [4].

Вот три основных типа ИМК:

- Неинвазивный ИМК, основанный на распознавании ментальных состояний, вызванных воображаемым выполнением движений. Он обеспечивает формирование дискретных управляющих команд и требует минимального времени обучения оператора при достаточно высокой производительности.
- Неинвазивный ИМК, использующий принцип непрерывного управления. После выработки навыка управления, такой ИМК позволяет управлять внешним устройством как собственным (виртуальным) исполнительным органом, не требуя ментального кодирования дис-

кретного набора команд.

- Инвазивный ИМК, основанный на двусторонней связи мозг-компьютер посредством имплантируемых электродов и позволяющий полностью инкорпорировать внешние технические устройства во внутреннюю нейронную модель схемы тела и, соответственно, оперировать с ними так же, как и с естественными исполнительными органами.

ИМК первого типа уже сейчас достигают информационной производительности 35 бит в минуту и требуют для обучения всего 20 минут. В свою очередь, ИМК второго типа, хотя и требуют длительного обучения испытуемых по биологической обратной связи, могут иметь и значительно более высокую производительность. В частности, информационная производительность так называемого водсвортского ИМК уже сейчас может достигать до 100 бит в минуту. К тому же, они могут обеспечивать возможность одновременного управления несколькими степенями свободы.

С точки зрения потенциальной производительности и способности управлять внешними устройствами с большим числом степеней свободы наиболее перспективными представляются инвазивные ИМК, основанные на преобразовании в команду для внешнего устройства активности большого числа отдельных нейронов. Это ИМК, разрабатываемые группой Николелиса. Уже в настоящее время их информационная производительность может достигать 100 бит в минуту, а число степеней свободы может превосходить два. Однако для их практического внедрения имеются существенные трудности: вживленные в мозг электроды зарастают соединительной тканью, что уменьшает время их эффективной работы, а открытое трепанационное отверстие, через которое проходит



соединительный кабель, является потенциальным источником опасности инфекционного заражения. Имеется высокий этический порог их использования, т.к. они требуют операционного вмешательства в мозг. Технология является очень дорогостоящей. Кроме того, их эффективность на настоящий момент сравнима с той, которую обеспечивают наилучшие неинвазивные ИМК.

Ближе всего к нашей задаче неинвазивные ИМК, основанные на анализе электроэнцефалограммы (ЭЭГ). Остановимся на них подробнее.

Первый в мире коммерчески доступный нейрокомпьютерный интерфейс представлен компанией Guger Technologies на выставке CeBIT 2010.

Устройство Intendix предназначено в первую очередь для парализованных пациентов больниц. Интерфейс готов к использованию через 10 минут настройки. После тренировки можно набирать тексты с экранной клавиатуры на скорости до 1,25 символа в секунду путём концентрации внимания на строках и столбцах, мерцающих поочерёдно. Шапка для электроэнцефалографии при этом снимает сигнал с коры головного мозга. Такая скорость является рекордной для прототипов подобных устройств.

Кроме набора символов, программное обеспечение Intendix умеет подавать различные сигналы по желанию пользователя, озвучивать текст, отправлять документы на печать и по электронной почте.

Стоимость прибора составляет около \$12 250.

NeuroSky — компания, базирующаяся в Сан-Хосе (Калифорнии, США). Основная технология — система «Нейро-компьютерный интерфейс».

Neurosky публично продемонстрировала две возможности, определяемые их датчиками - внимание и мышление. Эти два датчика независи-

мы, таким образом пользователь может например быть внимательным и задумчивым в одно и то же время. Другие датчики считывают беспокойство и сонливость. В отличие от конкурирующих продуктов, датчики не могут распознавать мысли или выражения лица, но зато подобные датчики продаются по намного меньшей цене.

NeuroSky создали игру для демонстрации своего продукта, которая позволяет игрокам двигать объектами, такими как автомобили или мебель, концентрируясь на них, и поднимать объекты, расслабляясь.

В июне 2008 года вышла система Neural Impulse Actuator компании OSZ Technology. Прибор NIA был введён на рынок как игровое приспособление и разработан в сотрудничестве с доктором Эндрю Юнкером (англ. Andrew Junker) из ВАТ, который впервые применил эту технологию в медицинских целях. Основным преимуществом NIA по сравнению со стандартными игровыми контроллерами и манипуляторами (компьютерная мышь, клавиатура, джойстик, геймпад) является существенное сокращение времени реакции от приблизительно 200 миллисекунд для компьютерной мыши до примерно 80 – 150 миллисекунд при использовании NIA.

NIA может улавливать три типа сигналов — микродвижения мышц лица (брови, уши, челюсть), движения глаз, альфа- (до 13 Гц) и бета- (от 14 Гц) волны мозга. По каждому из нескольких каналов (мышцы, глаза, мозговые волны) может проходить сигнал, поделенный на четыре уровня мощности по пороговому значению (threshold).

В 2009 году вышла в свет система Ерос компании Emotiv Systems. Она укомплектована 14 электродами (19 из них располагаются там же, где стандартные электроды при снятии ЭЭГ, 3 как в системе Neural Impulse Actuator, и 2 электрода как в системе NeuroSky). Также имеется двуххо-

севой гироскоп для измерения поворотов головы. Такая система обрабатывает 4 типа входов:

- сознательные мысли - 12 направлений движения, 6 направлений поворота и еще одну команду - "исчезнуть".
- эмоции - волнение, увлечение/скуку, медитацию, расстройство.
- выражения лица - позиции век и бровей, движения взгляда в горизонтальном направлении, улыбку, смех, стиснутые зубы, ухмылку. Эти данные считываются с мускулов лица и распознаются в течение 10 мс.
- повороты головы.

## 2.4 Метод наименьших квадратов

Классическим вариантом оценивания параметров по наблюдениям, осложненным шумами, является метод наименьших квадратов. Рассмотрим случай линейной функции одного аргумента. Пусть из опыта получены точки:

$$(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)$$

Требуется найти уравнение прямой

$$y = ax + b,$$

наилучшим образом согласующейся с опытными данными.

Пусть найдена такая прямая. Обозначим через  $\delta_i$  расстояние опытной точки от точки на этой прямой с тем же значением по оси абсцисс. В данном случае

$$y_i - ax_i - b = \delta_i$$

(3)

Чем меньше числа  $\delta_i$  по абсолютной величине, тем лучше подобрана прямая. В качестве характеристики точности подбора прямой можно принять сумму квадратов

$$S = \sum_{i=1}^n \delta_i^2$$

Покажем, как можно подобрать прямую таким образом, чтобы  $S$  была минимальной. В данном случае

$$S = \sum_{i=1}^n (y_i - ax_i - b)^2.$$

Условия минимума  $S$  задаются следующим образом:

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - ax_i - b) x_i = 0,$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n (y_i - ax_i - b) = 0.$$

Эти уравнения можно записать в таком виде:

$$\sum_{i=1}^n y_i x_i = a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i,$$

$$\sum_{i=1}^n y_i = a \sum_{i=1}^n x_i + nb.$$

Теперь можно легко найти  $a$  и  $b$  по опытным значениям  $x_i$  и  $y_i$ . Прямая,  $y = ax + b$ , называется прямой, полученной по методу наименьших квадратов (этим названием подчеркивается то, что сумма квадратов  $S$  минимальна).

Описанный выше метод легко обобщить на более сложные (нелинейные) зависимости [5].

## 3 Система «Окулографический интерфейс»

### 3.1 Предназначение и функциональность системы

Система является аналогом компьютерной «мыши», позволяя перемещать на мониторе курсор в соответствии с движениями глаз пользователя. Система предназначена в первую очередь для людей с ограниченными физическими возможностями. Но она может быть полезна всем, кто проводит много времени за компьютером.

Кроме того, система может выступать в качестве визуализатора электроокулограммы (ЭОГ). С помощью нее в НИИ Ухтомского проводят исследования зависимости различных явлений и параметров окуломоторной системы от пола, возраста и других особенностей человека.

Система имеет три режима работы: режим калибровки, режим обучающей игры и режим непосредственного управления компьютером. Опишем их подробнее.

*Калибровка* - настройки системы для конкретного пользователя-испытуемого. Человеку показывается стимул (закрашенный круг диаметром 5 мм), резко перемещающийся с интервалом в 1 секунду по часовой стрелке, и каждый раз возвращающийся в центр. Скачки идут по часовой стрелке, первый скачок происходит от центра вправо. Сначала малые скачки, потом средние, потом большие. На каждой окружности всего восемь положений стимула (см. рис. 2).

*Обучающая игра.* После проведения калибровок можно приступать к процедуре «обучающая игра». Испытуемому необходимо при помощи движений глаз поместить собственный курсор (синий круг) внутрь мишени - красного круга большего диаметра («накрыть» его). В это время экспериментатор может корректировать настройки. При «накры-

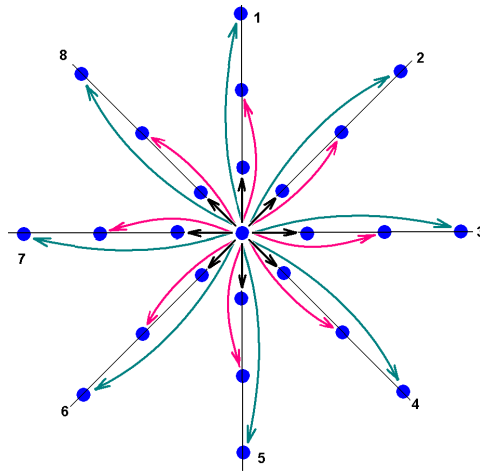


Рис. 2: Схема движения цели при калибровке

тии» мишени экспериментатор включает следующую мишень, либо испытуемый сам задает следующую мишень морганием правого глаза. Можно также поменять задание (не «накрывая» мишень) морганием левого глаза. Это позволяет тренировать не только навигацию курсора при помощи глаз, но и имитацию нажатий кнопок мыши при помощи монокулярных морганий, то есть морганий одним глазом. При достижении испытуемым определённого навыка есть возможность передать под его контроль и системный курсор.

Кроме того, с помощью системы возможен просмотр записей электроокулограмм и их преобразования.

### 3.2 Архитектура системы

На данный момент реализована экспериментальная установка с двумя мониторами (см. рис. 3). Для регистрации ЭОГ применяется 32-канальный электроэнцефалограф фирмы «Мицар» (Санкт-Петербург). При данной полосе пропускания усилителей (0.5 - 70 Гц) медленные

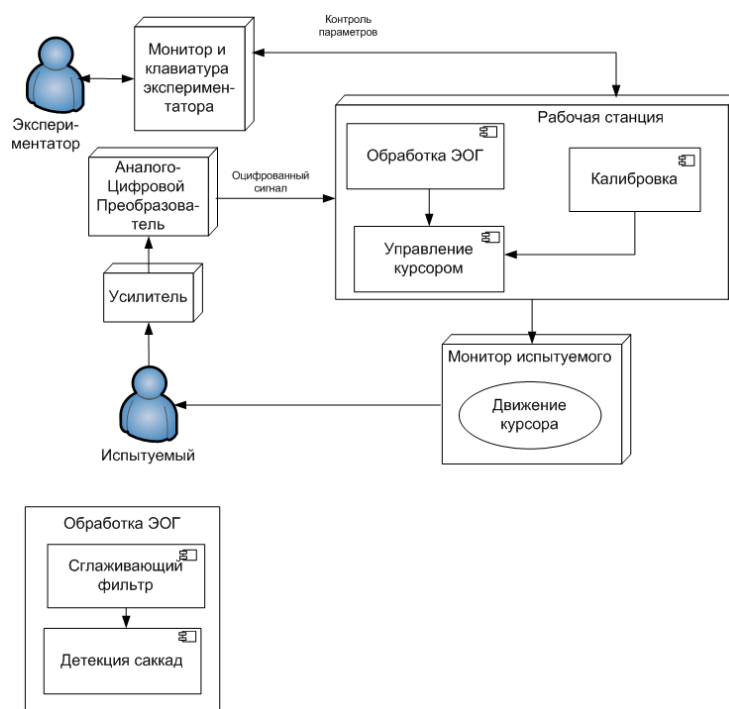


Рис. 3: Архитектура системы «Окулографический интерфейс»

смещения взгляда, характерные для визуального слежения или поворотов головы игнорируются автоматически, поскольку соответствуют частоте ниже 0.5 Гц. Блок усилителей укомплектован 32-канальным аналого-цифровым преобразователем (АЦП) фирмы L-Card (Москва). Частота дискретизации на выходе АЦП составляет 2000 Гц.

Программное обеспечение для системы разрабатывается в НИИ Ухтомского на протяжении 5 лет, с использованием среды разработки C++ Builder. Основная логика работы системы такова. Данные поступают малыми порциями и проходят через цепочку фильтров, параметры которых можно изменять динамически для корректировки работы экспериментатором. Данные хранятся в бинарных файлах в специально разработанном формате EEG. Система направлена в основном на предельную визуализацию и контроль в реальном времени. Также есть визуализатор,

при использовании которого не обязательно присутствие испытуемого и наличие АЦП и который может быть запущен на любом компьютере. Недостаток этого модуля в том, что он по-прежнему предоставляет максимальную визуализацию, но не предоставляет возможностей для сбора статистики. Поэтому статистика собирается вручную, правильность работы системы измеряется «на глаз» и по субъективным впечатлениям испытуемых.

При работе системы в режиме калибровки строятся калибровочные кривые, позволяющие по амплитуде саккад на каналах вычислить направление и количество пикселей, на которое должен сместиться курсор на мониторе испытуемого.

Блок обработки ЭОГ состоит из сглаживающего фильтра и компоненты детекции саккад.

*Сглаживающий фильтр* подавляет высокочастотные помехи. В момент времени  $i$  на вход поступает сигнал, пришедший с аналого-цифрового преобразователя -  $x_i$ . Выход сглаживающего фильтра  $S_i$  определяется по формуле

$$S_i = \alpha x_i + (1 - \alpha)S_{i-1},$$

где  $\alpha$  - параметр фильтра, а в качестве  $S_0$  берется усредненное значение сигнала на некотором начальном промежутке, на котором гарантированно нет движений глаз, то есть полезной составляющей. Каждый канал при данной фильтрации рассматривается отдельно.

*Компонента детекции саккад.* На вход принимается сглаженный сигнал (выход экспоненциального фильтра). Принцип работы фильтра таков: есть набор состояний и программа переключается между ними в зависимости от входных данных.



Таким образом, при поступлении новой порции ЭОГ-данных они обрабатываются и анализируются, при обнаружении саккад их амплитуды переводятся в смещение в пикселах, после чего курсор на мониторе испытуемого смещается.

### 3.3 Проблемы

Все пользователи-испытуемые, работавшие с «Окулографическим интерфейсом», сходятся во мнении, что управлять системой тяжело. Курсор часто двигается не в том направлении, что и глаза, с непропорциональной амплитудой, часто происходят ложные срабатывания. После анализа этих субъективных мнений был выделен ряд проблем, присущих системе.

Блок обработки ЭОГ работает не вполне приемлемо.

- Во-первых, он не всегда отличает передний фронт саккады от шума.
- Во-вторых, он дает ответ о начале движения только когда оно уже закончилось. Значит, и курсор переводится опозданием, а если еще учесть задержку, наведенную экспоненциальным фильтром, получается время ответа на действия пользователя 0.1 - 0.3 секунды, что существенно снижает удобство работы с системой.
- Кроме того, не в полной мере удастся использовать знания о физиологии саккадических движений, а именно об их согласованности по каналам. Так, если движение происходит, например, только в горизонтальном направлении, на каналах, отвечающих за вертикальное смещение, саккад не появится, поэтому ставятся довольно слабые условия на количество каналов, по которым саккада обязана быть

обнаружена. Это приводит к большому числу ложноположительных срабатываний.

### 3.4 Пути решения проблем

В ходе исследований физиологии окуломоторной системы человека обнаружено, что перед каждой саккадой у каждого человека регистрируются короткие негативные броски потенциала - *предсаккадные пики* (см. рис. 4).

Их предполагаемая природа такова. Перескок возбуждения с клетки на клетку сопровождается так называемым постсинаптическим потенциалом на мембране клетки-адресата (ПСП). Конкретно для перескока с нервного окончания на мышечную клетку этот потенциал имеет специальное название «потенциал концевой пластинки» (ПКП). Концевая пластинка - это контакт нерва и мышцы или нервно-мышечный синапс. Этот потенциал запускает цепь событий в мышечном волокне, результатом которых и является сокращение мышцы и, соответственно, движение глазного яблока. Но этот процесс занимает какое-то время и само сокращение (и движение) несколько отстаёт от ПКП, он уже успевает закончиться. Это мы и наблюдаем - предсаккадный пик уже закончился, а глазное яблоко только начало движение. Обычно ПСП и ПКП можно наблюдать только при помощи внедрённых в ткань игольчатых электродов. Однако глазные мышцы и глазодвигательные нервы имеют особые свойства по сравнению со скелетной мускулатурой, они очень мощные для своих размеров [6]. Поэтому ПКП виден даже при использовании накожных электродов. Его не видели раньше из-за шума или недостаточной частоты дискретизации (ПКП очень короткий - 10-15 мс), или из-за особенностей фильтрации - воспринимали пик как помеху, и избав-

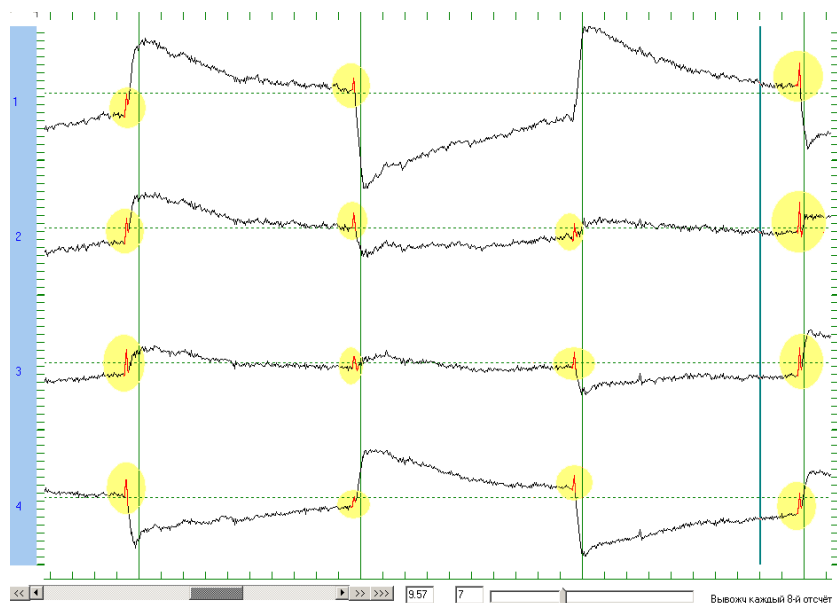


Рис. 4: Предсаккадные пики

лялись от него, и скорее всего на аппаратном уровне.

Использование информации о предсаккадных пиках поможет в полной мере учесть явление согласованности саккадических движений в системе «Окулографический интерфейс». Кроме того, это позволит сделать вывод о факте движения глаз еще до его начала. Это избавит от ложноположительных срабатываний детектора саккад.

Еще один путь решения проблем системы заключается в разбиении саккад на более мелкие фрагменты, чтобы система могла смещать курсор постепенно, и начать это делать, пока движение глаз еще не закончилось. Значит, необходимо создание дискретизирующего фильтра обнаружения саккад. Такой фильтр должен разбивать передний фронт саккады на несколько «ступенек» с тем, чтобы система интерпретировала их как несколько идущих друг за другом саккад и соответственно сократилось бы время отклика системы. Функцией дискретизирующего фильтра по-

мимо разбиения саккад может быть подавление низкочастотных помех.

Для снижения уровня высокочастотных помех, а значит, более точного определения амплитуды можно реализовать биполярную фильтрацию. Подробнее об этом см. дипломную работу Алексея Медведева (мат.-мех. ф-т СПбГУ, 545 гр.).

## 4 Техническая постановка задачи

Исходя из описанных выше особенностей системы «Окулографический интерфейс», обозначенных проблем и намеченных путей их решения, сформулированы следующие задачи.

- Изучить ЭОГ-метод слежения за взглядом и систему «Окулографический интерфейс».
- Формализовать явление предсаккадного пика, выявить и сформулировать требования к алгоритму его обнаружения.
- Разработать алгоритм обнаружения предсаккадных пиков в режиме реального времени с учетом их почти одновременного появления на всех каналах.
- Разработать программную реализацию алгоритма и интегрировать в систему «Окулографический интерфейс».
- Реализовать блок сохранения амплитуды саккады в рамках совместной с Алексеем Медведевым разработки дискретизирующего фильтра.

## 5 Предлагаемые алгоритмы

### 5.1 «Стихийный» алгоритм

Этот алгоритм практически не требует вычислительной мощности, его сложность  $O(1)$ . Его идея в том, чтобы сглаживанием подавить более высокие частоты, чем примерная частота пика, а потом отслеживать на сглаженном сигнале резкие скачки.

В качестве параметров алгоритм принимает полудлину предполагаемого пика  $l$ , пороговое значение разницы между значениями в начальной и конечной точках  $d$ , пороговое значение доли от общего количества каналов, на которых обнаружен пик  $q$  и значение допустимой временной разницы между одним и тем же пиком, детектированным на разных каналах  $p$ . Здесь и далее  $N$  - это количество включенных каналов. Обозначим порцию сигнала (отсчет), поступающего с АЦП в некоторый промежуток времени, через  $X = x_1, \dots, x_N$ . Выход алгоритма - число  $a$ , равное 1, если по мнению алгоритма на данном отсчете пик присутствует, и 0 в противном случае. Для хранения промежуточных данных используется двумерный кольцевой буфер  $s$  с размерностями  $l$  и  $N$ , в который записываются последние  $l$  значений входного сигнала на  $N$  каналах, и одномерный буфер  $k$  размерностью  $N$ , в котором хранится информация о том, сколько времени прошло с момента последней детекции пика на данном канале.

Работает алгоритм следующим образом. В начале буфер заполняется точками поступившего сигнала:

$$s_{ir} = x_i$$

$$r = (r + 1) \text{ mod } l$$

Далее, для каждого канала происходит анализ того, имеется скачок сигнала или нет, и информация заносится в буфер  $k$ : if  $|x_i - s_{ir}| \geq d$  then  $k_i = p$  else  $k_i = \max(k_i - 1, 0)$ .

Далее, если выполняется условие

$$\frac{\sum_{i=1}^N \frac{k_i}{|k_i|}}{N} \geq q, \quad (1)$$

где сумма берется по каналам, на которых  $k_i \neq 0$ , то делается вывод о наличии пика:  $a = 1$ . Если же это условие не выполняется, то возвращается  $a = 0$ .

Этот алгоритм в ходе апробации показал слишком большое количество ложноположительных срабатываний, поэтому встала задача разработать более эффективный вариант.

## 5.2 Адаптированный метод наименьших квадратов

Данный алгоритм позволяет сравнивать анализируемый сигнал с любым наперед заданным шаблоном. Он заключается в следующем.

Пусть длина шаблона  $l$ , сам шаблон представлен в виде массива точек  $y[1..l]$ . Нужно найти участки исходного сигнала  $x$ , которые похожи на данный шаблон. Для каждого участка длины  $l$  найдем коэффициенты  $\alpha$  и  $\beta$ , такие, что сигнал  $x_i$  максимально похож на  $\alpha y_i + \beta$ . Будем искать такие промежутки, на которых сигнал может двигаться по абсолютному значению (все вместе), но не друг относительно друга. Также допускается растяжение-сжатие шаблона в некоторых пределах. Для этого нужно найти такие  $\alpha$  и  $\beta$ , чтобы выражение

$$\sum_{i=1}^l (\alpha y_i + \beta - x_i)^2$$

было минимально, и оценить сумму квадратов (сравнить с пороговым значением). Кроме всего перечисленного, отслеживается параметр  $\alpha$ . При поступлении новой точки следует сравнить последние известные  $l$  точек сигнала с шаблоном. Для этого берем производные по  $\alpha$  и по  $\beta$  и приравняем нулю.

$$\begin{cases} \sum_{t=1}^l x_i((\alpha - 1)x_i + \beta) = 0 \\ \sum_{t=1}^l ((\alpha - 1)x_i + \beta) = 0 \end{cases}$$

Из этой системы следует, что если

$$\alpha = \frac{\sum_{t=1}^l s_{it} \cdot y_t - \frac{\sum_{t=1}^l s_{it} \sum_{t=1}^l y_t}{l}}{\sum_{t=1}^l y_t^2 - \frac{\left(\sum_{t=1}^l y_t\right)^2}{l}}, \quad (2)$$

$$\beta = \frac{\sum_{i=1}^l x_i - \alpha \sum_{i=1}^l y_i}{l}, \quad (3)$$

то можно считать, что данный участок в достаточной мере похож на шаблон с учетом параметра  $\alpha$ .

В качестве параметров алгоритм принимает массив-шаблон  $y[1..l]$ , пороговое значение для суммы квадратов разностей  $z$  и порог  $d$  для получившегося  $\alpha$ , а также пороговое значение доля от количества каналов, на которых должен быть обнаружен пик  $q$  и значение допустимой разницы во времени между обнаружением пика на разным каналах  $p$ . Пусть  $N$  - количество включенных каналов, а  $X = x_1, \dots, x_N$  - это вектор сигнала, поступающего в каждый момент времени. Выход алгоритма - ответ на вопрос, есть ли пик на данном отсчете. Для хранения промежуточных данных используется двумерный кольцевой буфер  $s$  с размерностями  $l$



и  $N$ , в который записываются последние  $l$  значений входного сигнала на  $N$  каналах, и одномерный буфер  $k$  размерностью  $N$ , в котором хранится информация о том, сколько времени прошло с момента последней детекции пика на соответствующем канале. Работает алгоритм следующим образом:

$$s_{ir} = x_i$$

$$r = (r + 1) \bmod l$$

Для каждого канала по формулам 2 и 3 вычисляются  $\alpha$  и  $\beta$ . Если выполняются неравенства

$$\alpha \geq d$$

и

$$\sum_{i=1}^l (\alpha y_i + \beta - x_i)^2 \leq zl,$$

то

$$k_i = p,$$

иначе  $k_i = \max(k_i - 1, 0)$

Далее, если выполняется условие

$$\frac{\sum_{i=1}^N \frac{k_i}{|k_i|}}{N} \geq q, \quad (4)$$

(сумма берется по каналам, на которых  $k_i \neq 0$ ), то делается вывод о наличии пика:  $b = 1$ , иначе констатируется его отсутствие на данном отсчете:  $b = 0$ .

Ограничение высоты снизу необходимо для того, чтобы отсеять участки, для которых коэффициент  $\alpha$  отрицателен, то есть зеркально отраженные, а также с  $\alpha$ , близким к нулю, что соответствует почти постоянным участкам.

Кроме того, контроль высоты участков оказался весьма полезен для проверки гипотезы о том, что пики имеют одинаковую высоту для одного человека, причем для каждого человека - свою, в ходе апробации компоненты обнаружения предсаккадных пиков в тестовом окружении.

Данный алгоритм тестировался с шаблоном-синусоидой на сглаженном сигнале, с л-образным на несглаженном сигнале и с одним из пиков определенного человека в качестве шаблона также на необработанном сигнале.

Сложность алгоритма  $O(l)$ , где  $l$  - длина шаблона.

## 6 Реализация

### 6.1 Функциональность и интеграция в систему

*Язык и средства разработки.* Дискретизирующий фильтр и компонента обнаружения предсаккадных пиков реализованы на языке C++ в среде C++ Builder 6. Выбор языка и среды обусловлен необходимостью соответствия средствам реализации системы «Окулографический интерфейс».

*Функциональность* компоненты обнаружения предсаккадных пиков заключается в следующем: в каждый момент времени на вход поступают значения ЭОГ со всех каналов, на выходе компонента выдает ответ на вопрос, есть ли предсаккадный пик на данном отсчете.

Функциональность дискретизирующего фильтра состоит в разбиении детектируемых саккад на меньшие части и подавлении низкочастотных помех. В частности, блок сохранения амплитуды саккады следит за тем, чтобы суммарная амплитуда была равна амплитуде саккады, которая бы детектировалась без разбиения.

*Интеграция в систему.* Диаграмма, отражающая архитектуру системы «Окулографический интерфейс», представлена на рис. 3. Компонента обнаружения предсаккадных пиков встраивается в блок «Обработка ЭОГ» (см. рис. 5).

Дискретизирующий фильтр занимает место после сглаживающего фильтра.

Между компонентой обнаружения предсаккадных пиков и дискретизирующим фильтром организовано совещание. Результат дискретизирующего фильтра учитывается только если с момента последнего положительного ответа от компоненты обнаружения пиков прошло не очень

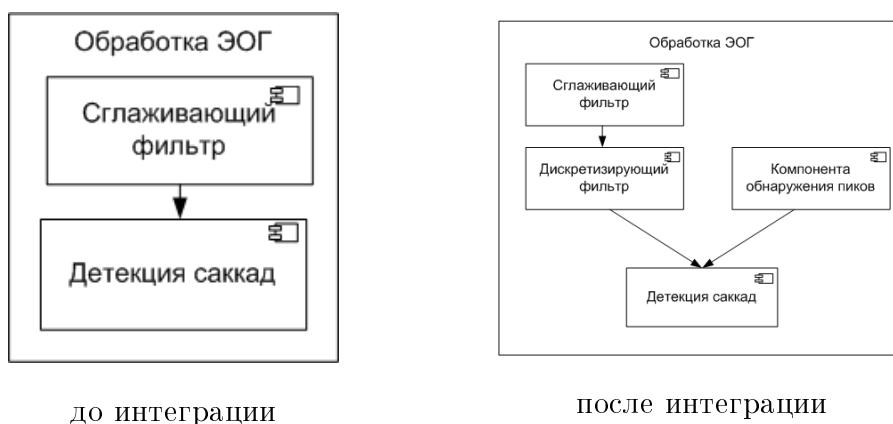


Рис. 5: Блок «Обработка ЭОГ»

много времени (не больше порогового значения, выступающего в роли параметра системы).

Это изменило и компоненту «Детекция саккад», которая теперь выполняет функции синхронизации результатов дискретизирующего фильтра и компоненты обнаружения предсаккадных пиков.

## 6.2 Особенности реализации

На электроокулограмме не очевидно местоположение саккад, а тем более предсаккадных пиков, которые визуалью трудно отличимы от высокочастотных помех. Помимо этого, сложно формализовать требования к самому пику, потому что при визуальном определении экспертом используется информация о наличие или отсутствии идущей за пиком саккады, для алгоритма же обнаружения предсаккадных пиков такая информация недоступна, потому что он работает отдельно от детекции саккад и представляет собой независимый критерий обнаружения факта движения глаз.

Система «Окулографический интерфейс» разрабатывалась в основном профессиональным биологом, она довольно хорошо спроектирована

и продумана, но реализована с использованием устаревших концепций и не принятых в современной программной инженерии методов, а также плохо документирована. Поэтому возникали определенные сложности в понимании архитектуры системы. Кроме того, использовался бинарный формат хранения электроокулограмм (формат .eeg) и, соответственно, низкоуровневые небезопасные функции работы с ними. В этом формате хранилось большое количество избыточной информации (например, номера начальных байтов фрагментов записи). С бинарным форматом данных также связана сложность тестирования и отладки, потому что совершенно неочевидно, как отрезать кусочек для тестирования от длинной (примерно 200000 отсчетов) записи. В связи с этим был создан конвертор, преобразующий записи электроокулограмм из бинарного формата в текстовый. Сложности тестирования заключается еще и в том, что если оставаться в рамках существующего окружения, то невозможно получить никакие параметры эффективности работы компоненты, кроме визуальных. Это привело к необходимости создания независимого тестового окружения и методики тестирования для компоненты обнаружения предсакдных пиков.

Далее, большое неудобство при интеграции заключалось в том, что система не работает без специального оборудования (платы АЦП), которое есть только в НИИ Ухтомского.

Отдельно о разработке дискретизирующего фильтра следует отметить, что она велась совместно с Медведевым Алексеем, что повлекло необходимость разделения задачи на мелкие части необходимости их дальнейшей интеграции.

## 7 Тестовое окружение, тестирование и апробация

Ниже пойдет речь о тестировании и апробации компоненты обнаружения предсаккадных пиков.

### 7.1 Архитектура и средства реализации

Тестовое окружение для компоненты обнаружения предсаккадных пиков состоит из конвертора, обработчика и визуализатора (см. рис 6).

*Конвертор* переводит электроокулографические данные из бинарного формата в текстовый, содержащий только те сведения, которые необходимы для данной задачи. Формат выходных данных следующий: номер отсчета, вектор ЭОГ, вектор меток. Размерность вектора электроокулограммы равна количеству каналов, размерность вектора меток задается при записи электроокулограммы в файл системой «Окулографический интерфейс». Разные фрагменты одной записи переводятся в разные файлы для удобства дальнейшего использования.

*Обработчик* по заданной (загруженной из конфигурационного файла) конфигурации вызывает компоненту обнаружения предсаккадных пиков с разными значениями параметров, каждый раз считая значения критериев качества работы компоненты.

*Визуализатор* получает массив значений изменявшегося параметра и массивы соответствующих значений критериев качества, и по этим данным строит графики.

*Средства реализации.* Конвертор реализован на языке C++, в среде C++ Builder 6 в целях максимального переиспользования имеющегося кода (описание бинарного формата). Обработчик и визуализатор реали-

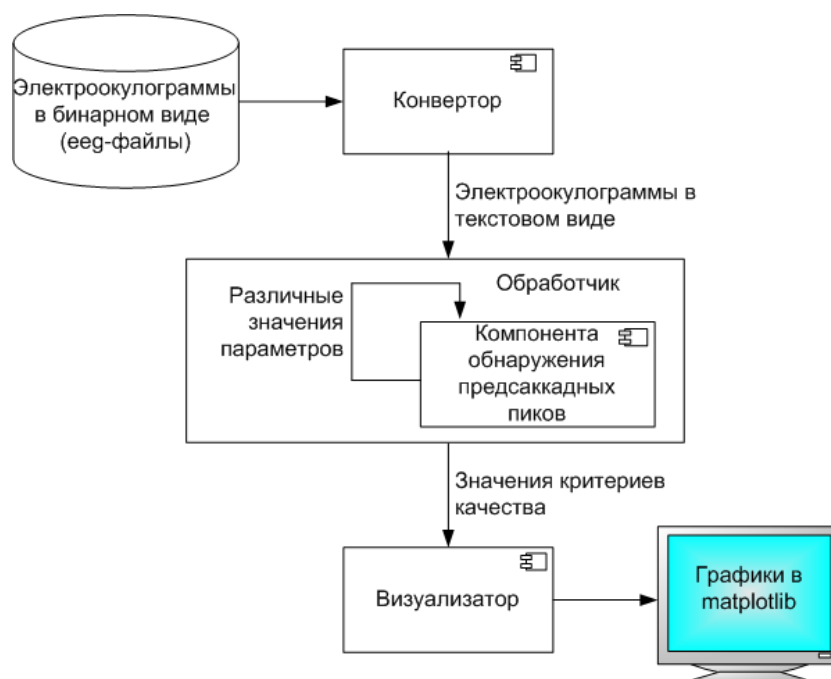


Рис. 6: Тестовое окружение

зованы на языке Python (версия 2.6), последний реализован с использованием библиотеки matplotlib.

## 7.2 Методика тестирования

В качестве тестовых данных брались записи разных людей, полученные при работе системы «Окулографический интерфейс» в режиме калибровки. Как правило, при работе в таком режиме заказывается два канала меток, первое значение показывает, куда в данный момент двигался стимул (если он двигался), а второе - отмечен ли этот отсчет при просмотре записи «вручную». На записях предсаккадные пики отмечались «вручную», использовалась информация о том, что глаза начинают двигаться через 0.1 - 0.3 секунды после движения стимула.

При определении эффективности работы учитывается возможность

обнаружения пика с некоторым опозданием.

На вход обработчику поступают два вектора: вектор электроокулограммы  $S = (s_1, s_2, \dots, s_n)$  и вектор отмеченных пиков  $M = (m_1, m_2, \dots, m_n)$ , где  $n$  - количество отсчетов в анализируемой записи. Также поступает параметр  $t_s$  - на сколько отсчетов можно опоздать с сигналом о новом пике.

Сначала из вектора отмеченных пиков конструируем вектор оценок  $K = (k_1, k_2, \dots, k_n)$  следующим образом:

$$k_i = \begin{cases} 1 & \text{если } m_i = 1 \\ (k-1)^2 & \text{если } k < t_s \text{ где } k = i - l, l = \max(i_t) : i_t < i, m_{i_t} = 1 \\ 0 & \text{иначе} \end{cases}^3$$

Это делается, потому что пики не равны по ширине, а следовательно, возможна ситуация, когда детектор даст положительный ответ о наличие пика в точке, которая будет следовать за отмеченными, но сама отмечена не будет, если пик будет уже шаблона. То есть детектор пиков сработает с опозданием. Это не является ошибочным срабатыванием. Поэтому таким точкам будут соответствовать числа  $k : 0 \leq k \leq 1$ , убывающие с удалением от конца отмеченного пика. Такой прием оказался нужным в основном при использовании «стихийного» алгоритма из-за запаздывания сигнала после сглаживания. В случае метода наименьших квадратов все проблемы удастся решить правильным выбором параметров шаблона. Ниже описаны критерии оценки качества обнаружения пиков.

- *Доля обнаруженных пиков.* Каждый пик рассматривается отдельно и помечается найденным, если хотя бы раз за его протяженность детектор пиков дал положительный ответ. Количество помеченных пиков делится на общее количество пиков.



- *Количество ложных срабатываний.* Считается количество ситуаций, в которых компонента обнаружения предсаккадных пиков сообщила о наличие пика, но соответствующая компонента вектора оценок равна нулю. Полученное число делится на количество отсчетов в секунде (в тестовых примерах это 2000).
- *Полезная производительность.*

$$h_i = m_i \cdot s_i,$$

$$Prod = \frac{\sum_{i=1}^n h_i}{M},$$

где  $M$  - количество ненулевых слагаемых в числителе. Эта оценка отражает то, насколько точно детектор распознает пики, то есть велик ли процент опозданий. Она ничего не говорит об ошибочных срабатываниях. Если полезная производительность равна 1, это означает, что все пики обнаружены «вовремя», чем ближе полезная производительность к 0, тем больше произошло опозданий. Опять же, стоит отметить, что этот критерий значим только для «стихийного» алгоритма.

Таким образом, методика тестирования компоненты обнаружения предсаккадных пиков заключается в следующем. Тестовые данные обрабатываются компонентой обнаружения предсаккадных пиков, по ее ответам и меткам считаются значения критериев качества, что в случае тестирования позволяет найти ошибки в реализации, сравнивая результаты с предполагаемыми.

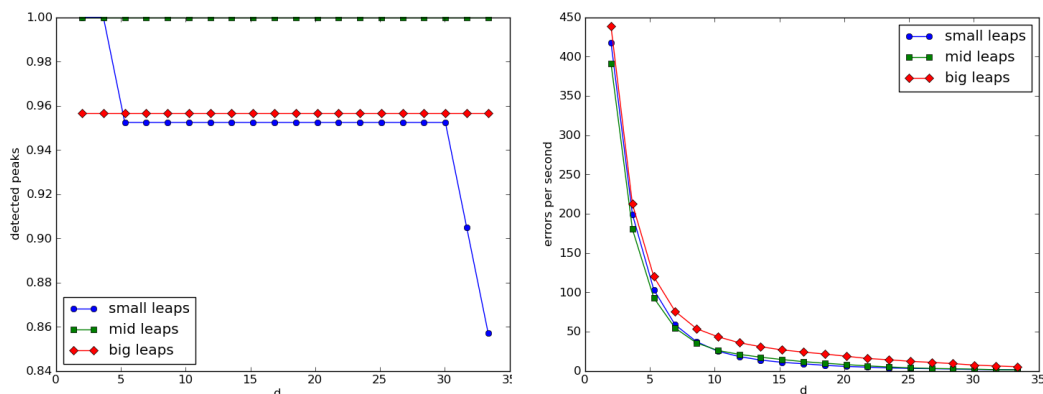
Кроме того, данное тестовое окружение применяется для апробации компоненты обнаружения предсаккадных пиков. В этом случае данные обрабатываются не с целью проверки правильности работы компоненты,

а с целью ее настройки (подбора параметров) для конкретного пользователя, или выявления зависимости параметров пиков от пола, возраста и индивидуальных особенностей пользователей-испытуемых. То есть тестовое окружение выступает в роли инструмента для физиологических исследований, и может быть использован специалистами в предметной области.

## 8 Результаты апробации

Для апробации брались записи калибровки, размеченные «вручную». Отдельно строились кривые для больших, средних и малых скачков взора.

«Стихийный» алгоритм. На рис. 7 приведены графики эффективности работы «стихийного» алгоритма (см. раздел 5.1) от параметра  $d$  (при этом  $l = 5$ ,  $q = 0.5$ ,  $p = 1$ , параметр предварительного сглаживания  $\alpha = 0.06$ ).



(а)

(б)

Рис. 7: Зависимость эффективности обнаружения пиков «стихийным» алгоритмом от параметра  $d$ : (а) - доля обнаруженных пиков, (б) - количество ложных срабатываний в секунду

Как и ожидалось, процент ошибок падает при увеличении желаемой амплитуды, то есть параметра  $d$ , потому что амплитуда ограничивается снизу, а значит, при увеличении порога остается все меньше участков с подходящими амплитудами. Процент детектируемых пиков приемлем при  $d \leq 30$ . Однако количество ложных срабатываний совершенно неприемлемо. Это говорит о незрелости алгоритма и необходимости при-

менения других методов к задаче обнаружения предсаккадных пиков.

*Компонента обнаружения предсаккадных пиков*, основанная на адаптированном методе наименьших квадратов, исследовалась на неотфильтрованных данных с л-образным шаблоном и с одним из пиков в качестве шаблона.

Она оказалась значительно более эффективной, чем та, которая использовала «стихийный» алгоритм. На рис. 8 приведены графики зависимости количества ложных срабатываний в секунду и доли обнаруженных пиков от высоты л-образного шаблона в случае больших, средних и малых по амплитуде саккад. При этом  $q = 0.8$ ,  $p = 1$ ,  $d = 0.9$ , полудлина  $l = 10$  (см. рис. 8).

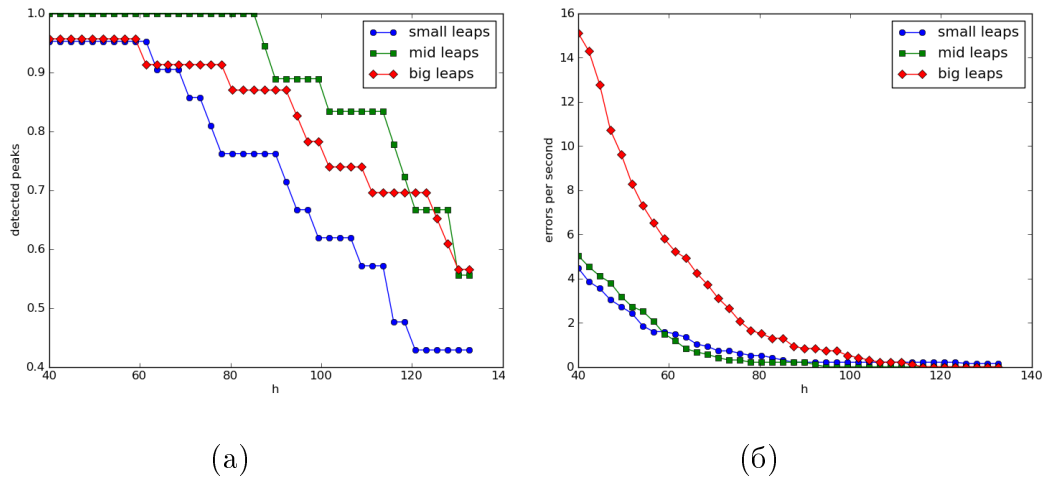


Рис. 8: Зависимость эффективности детекции пиков от частоты шаблона: (а) - доля обнаруженных пиков, (б) - количество ложных срабатываний в секунду

По сравнению со «стихийным» алгоритмом количество ошибок заметно уменьшилось. Это говорит о большей эффективности данного подхода.

Далее, исследовалась зависимость эффективности работы компоненты от полудлины л-образного шаблона  $l$  при  $q = 0.8$ ,  $p = 1$ ,  $d = 0.9$ , полудлина  $h = 70$  (см. рис. 9).

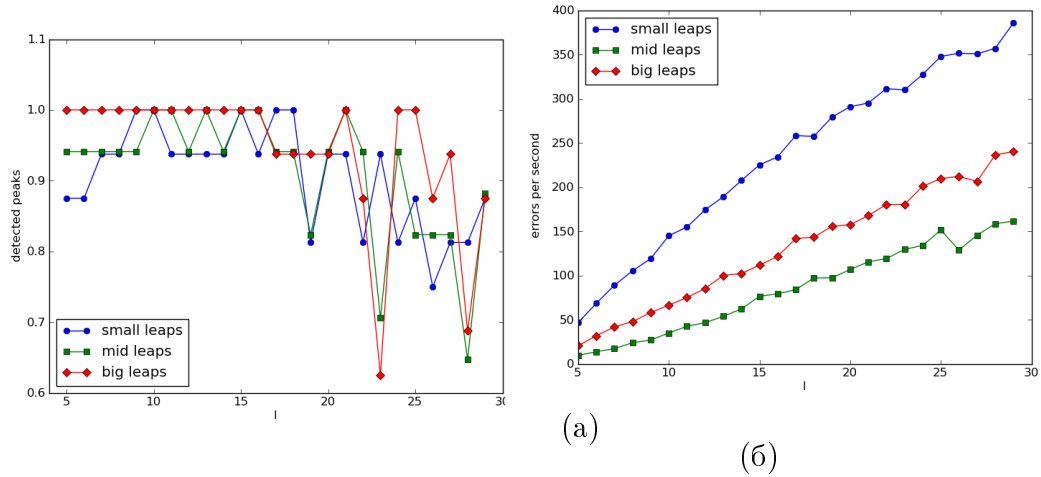


Рис. 9: Зависимость эффективности детекции пиков от параметра  $l$ : (а) - доля обнаруженных пиков, (б) - количество ложных срабатываний в секунду

Исходя из этого графика, можно судить о том, что при данных параметрах пики оптимально детектируются при полудлине шаблона от 10 до 15 отсчетов. Такой разброс длины объясняется тем, что с шаблоном может успешно сравниваться не только целый пик от начала до конца, но, возможно, его некоторая часть. При увеличении параметра  $q$  вдвое снижается количество ошибок, но несколько уменьшается и доля правильно обнаруженных пиков. Так что, если не обязательно распознавать все пики, можно, подобрав по-другому параметры, существенно снизить процент ошибки.

Отметим, что, как и предполагалось, параметры пиков, а значит, и оптимальные параметры фильтров, для разных людей - разные. До сих

пор были приведены зависимости, характеризующие одного человека и подобраны оптимальные параметры работы для него:  $q = 0.8$ ,  $p = 1$ ,  $d = 0.9$ , полудлина  $l = 10$ ,  $h = 70$ .

При апробации на записях калибровок другого человека найденное оптимальное значение параметра  $h$  было равно 100.

При использовании компоненты обнаружения предсаккадных пиков с одним из пиков в качестве шаблона можно полностью исключить ложноотрицательные срабатывания и свести количество ложноположительных срабатываний до 2-3 в секунду (то есть в 2000 отсчетов) на пяти электродах при  $q = 0.9$ ,  $p = 2$ ,  $d = 0.3$  на неотфильтрованном пятиканальном сигнале.

Преимущество такого шаблона в том, что не нужно строить математическую модель пика.

Вообще говоря, процент ошибочных срабатываний велик, объяснение этому следующее. В ходе исследований предсаккадных пиков с помощью компоненты обнаружения пиков было замечено, что не всякий такой пик предшествует саккаде с ощутимой амплитудой. Были замечены пики, по форме ничем не отличающиеся от предсаккадных, но стоящие обособленно. Эксперты в предметной области объясняют это явление следующим образом. Даже во время периодов фиксации взгляда, которые длятся примерно от 0,15 до 2 сек, глаза немного смещаются за счет глазного тремора, который никогда не прекращается (его амплитуда - несколько угловых минут, а частота находится в пределах 20-150 Гц). Во время продолжительной произвольной фиксации глаз происходит также медленный «дрейф», за счет которого точка фиксации уходит от фиксируемого объекта (в данном случае стимула - закрасенного круга). Компенсация этих дрейфов и восстановление правильной фиксации объекта проис-

ходит с помощью произвольных коррекционных микросаккад. Перед ними, как и перед любыми саккадами, появляются предсаккадные пики. Таким образом, кроме стандартных предсаккадных пиков компонентой детектируются пики коррекционных саккад. Возможно частично убрать коррекционные пики за счет повышения высоты шаблона. Все же они слишком похожи на пики перед обычными саккадами, поэтому необходимо организовывать совещание компоненты обнаружения предсаккадных пиков с дискретизирующим фильтром.

Кроме того, исходя из графиков, можно сделать вывод о том, что пики, относящиеся к одному человеку, имеют разную высоту.

И последнее. Из приведенных графиков видно, что амплитуда пиков не коррелирует с амплитудой саккады. Исследователи-биологи приводят такое объяснение этому факту. Источник тока для пика - мембраны мышечных клеток (постсинаптический потенциал). Поэтому амплитуда пика зависит просто от расстояния от электрода до мышцы, которое при поворотах глазного яблока не меняется или меняется мало. Из этого следует, что амплитуда на данном канале примерно постоянна и определяется в основном положением электрода. Амплитуда пика может лишь немного изменяться при смене направления (но не сильно) в зависимости от того, какой именно из шести мышц предстоит дернуться сильнее - более близкой к электроду или более дальней. Соответственно «вклады» отдельных мышц меняются, а ближняя мышца имеет больший вес. Амплитуда же самой саккады зависит от угла поворота глазного диполя и очень сильно зависит от направления.

## 9 Заключение

В рамках данной дипломной работы достигнуты следующие результаты.

- Были собраны и проанализированы требования к алгоритму обнаружения предсаккадных пиков.
- Произведено знакомство с архитектурой и реализацией системы «Окулографический интерфейс», а именно, с бинарным форматом представления ЭОГ-данных и с блоком обработки ЭОГ.
- Разработан алгоритм для обнаружения предсаккадных пиков (первая, «стихийная», версия и алгоритм на основе адаптированного метода наименьших квадратов).
- Алгоритм реализован и интегрирован в систему «Окулографический интерфейс» (C++, C++ Builder).
- Создано тестовое окружение для компоненты обнаружения предсаккадных пиков (язык Python). В рамках решения этой задачи разработана методика тестирования алгоритма, реализован конвертор данных электроокулограммы из бинарного формата в текстовый (C++, C++ Builder) и визуализатор статистики (язык Python, библиотека matplotlib).
- С помощью тестового окружения выполнен пример подбора параметров для оптимальной работы алгоритма обнаружения предсаккадных пиков для конкретного человека.
- Выполнена реализация блока сохранения амплитуды в рамках разработки дискретизирующего фильтра.



Эти результаты были успешно применены коллективом НИИ Ухтомского для исследования окуломоторной системы человека. В частности, в ходе апробации компоненты обнаружения пиков замечены коррекционные микросаккады, после чего проанализирована их природа. По графикам, созданным с помощью визуализатора, найдены зависимости параметров пиков от пола, возраста, индивидуальных особенностей испытуемого.

В качестве дальнейшей работы предполагается реализовать автоматический подбор параметров фильтров по записям калибровки для конкретного человека, развить подсистему обработки статистики, дополнить ее пользовательским интерфейсом для удобства работы исследователей-биологов. Кроме того, планируется расширить тестовое окружение для оценки качества работы всей системы (на данный момент тестовое окружение реализовано лишь для компоненты обнаружения предсаккадных пиков).

## Список литературы

- [1] Биология и медицина. <http://medbiol.ru>.
- [2] А. А. Демидов, А. В. Жегалло. Оборудование СМТ для регистрации движений глаз: тест-драйв. *Экспериментальная психология*, 1:149–159, 2008.
- [3] И. А. Секерина. Экспериментальные методы в психолингвистике. *Московский семинар по когнитивной науке*, Апрель 2005.
- [4] А. А. Фролов, В. Ю. Рощин. Интерфейс мозг-компьютер. Реальность и перспективы. *Научная конференция МИФИ*, 2008.
- [5] Я. Б. Шор. *Статистические методы анализа и контроля качества и надежности*. М.:Госэнергоиздат, 1962.
- [6] С. Ю. Эрам. Детекция движений глаз по электроокулограмме при смещениях взора. Магистерская диссертация, СПбГУ, 2007.
- [7] J. Anders. Pilot's attention allocation during approach and landing - eye- and head-tracking research in an a330 full flight simulator. In *International Symposium on Aviation Psychology (ISAP)*, 2001.
- [8] Arrington research. eye tracking systems, 2009. <http://www.arringtonresearch.com>
- [9] D. P. Belov, S. Yu. Eram, S. F. Kolodyazhnyi, I. E. Kanunikov and O. V. Getmanenko. Electrooculogram Detection of Eye Movements on Gaze Displacement. *Neuroscience and Behavioral Physiology*, 40(5):583–591, 13 May 2010.

- [10] Andreas Bulling, Daniel Roggen, and Gerhard Tröster. Wearable eog goggles: Seamless sensing and context-awareness in everyday environments. *Journal of Ambient Intelligence and Smart Environments*, 1(2):157–171, 4 2009.
- [11] A. T. Dutchowski. A breadth-first survey of eye tracking applications. *Behavior Research Methods, Instruments and Computers (BRMIC)*, 34(4):455–470, 2002.
- [12] Eye-tracking (images). <http://www.eyewriter.org/images/TEMP-T-ONE/eye-tracking/>
- [13] M. A. Frens and J. N. van der Geest. Recording eye movements with video-oculography and scleral search coils: a direct comparison of two methods. *Journal of neuroscience methods*, 114:185–195, 2002.
- [14] M. Land, N. Mennie, and J. Rusted. The role of vision and eye movements in the control of activities of daily living. *Perception*, 28(11):1307–1432, 1999.
- [15] J. Ozyurt, P. DeSouza, P. West, R. Rutschmann, and M.W. Greenlee. Comparison of cortical activity and oculomotor performance in the gap and step paradigms. In *European Conference of Visual Perception (ECVP)*, 2001.
- [16] J. Raphael. Google eye-tracking watches what you’re watching, February 2009.
- [17] D.D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Eye Tracking Research and Application Symposium*, page 71–78. ACM Press, 2000.

- [18] J. B. J. Smeets and I. T. C. Hooge. Nature of variability in saccades. *Journal of Neurophysiology*, 90:12–20, February 2003.
- [19] D. M. Snodderly, I. Kagan, and M. Gur. Selective activation of visual cortex neurons by fixational eye movements: Implication of neural coding. *Visual Neuroscience*, 18:259–277, 2001.
- [20] V. Tanriverdi and R. J. K. Jakob. Interacting with eye movements in visual environments. In *Human Factors and Computing Systems*, pages 265–272. ACM Press, 2000.