

ФИЗИЧЕСКИЕ ПРИБОРЫ ДЛЯ ЭКОЛОГИИ,
МЕДИЦИНЫ, БИОЛОГИИ

УДК 519.6:004.896

СРАВНЕНИЕ МЕТОДОВ ФИЛЬТРАЦИИ ДАННЫХ, ПОЛУЧЕННЫХ
С ПОМОЩЬЮ РАЗЛИЧНЫХ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ,
ПРЕДНАЗНАЧЕННЫХ ДЛЯ ДЕТЕКТИРОВАНИЯ ГОЛОВЫ
ПАЦИЕНТА ПРИ ЗАПИСИ vНIT-ТЕСТА НА КАМЕРУ СМАРТФОНА

© 2026 г. Г. Д. Мамыкин^{a,*}, А. А. Кулеш^b, Ф. Л. Барков^c,
В. П. Первадчук^a, А. О. Осколкова^b, С. Крапивин^b,
Ю. А. Константинов^c

^a Пермский национальный исследовательский политехнический университет
Россия, 614990, Пермь, Комсомольский просп., 29

^b Пермский государственный медицинский университет им. академика Е.А. Вагнера
Россия, 614000, Пермь, ул. Петропавловская, 26

^c Институт механики сплошных сред
Пермского федерального исследовательского центра Уральского отделения
Российской академии наук

Россия, 614013, Пермь, ул. Академика Королева, 1

*e-mail: matyckin.gennady@yandex.com

Поступила в редакцию 29.07.2025 г.

После доработки 17.09.2025 г.

Принята к публикации 30.09.2025 г.

Рассматривается задача повышения точности определения положения головы пациента при проведении медицинского теста vНIT с использованием видеозаписи, полученной с камеры смартфона. Основное внимание уделяется сравнению выходных данных трех моделей машинного обучения (Dense Head Pose Estimation, Head Pose Estimation и Lightweight Head Pose Estimation), предназначенных для детектирования положения головы, с эталонными данными, полученными с помощью IMU-сенсоров. Для повышения точности сигналов, полученных из видеопотока, было проведено сравнение четырех методов цифровой фильтрации, таких как эмпирическая модовая декомпозиция (EMD), вариационная модовая декомпозиция (VMD), дискретное вейвлет-преобразование и фильтрация Баттервортса. В качестве метрики качества использовалась среднеквадратичная ошибка (MSE), взвешенная по амплитуде сигнала. Результаты показали, что наилучшее соответствие эталонным данным достигается при использовании модели LHPE совместно с вейвлет-преобразованием и фильтрацией Баттервортса. Представленный подход подтверждает возможность

использования камеры смартфона для детектирования головы с высокой точностью, что открывает перспективы использования обычных смартфонов в качестве аппаратной части теста vHIT.

1. ВВЕДЕНИЕ

В настоящее время при обращении пациента с головокружениями одной из основных процедур для постановки диагноза является тест vHIT (video head impulse test). Это клинический тест, который оценивает адекватность компенсаторных движений глаз в ответ на внезапный поворот головы малой амплитуды в плоскости полукружных каналов. Цель тестирования – выявить нарушение вестибуло-окулярного рефлекса (VOR) у пациентов с головокружением, особенно при подозрении на острую периферическую вестибулопатию. Проблема диагностики причин головокружений у пациента с помощью vHIT актуальна в силу того, что используемое профессиональное оборудование является чрезвычайно дорогостоящим. Существует необходимость в разработке решения – как аппаратного, так и алгоритмического, которое будет дешевле имеющихся продуктов и при этом будет не сильно уступать им по качеству результатов.

Для того чтобы понять, в каких аспектах можно сократить стоимость конечного решения, нужно разобраться в текущих устройствах: проанализировать их аппаратную и алгоритмическую части. Ясно, что производители таких устройств не публикуют в открытом доступе все подробности их решений, однако по общим характеристикам уже можно строить гипотезы или предположения по архитектуре будущего улучшенного решения. Предполагается, что в качестве устройств для записи vHIT можно использовать заднюю/переднюю камеру современного смартфона. Таким образом, изучив текущие решения, мы сможем составить набор ограничений и требований к аппаратной части, а также составить методологию проведения эксперимента vHIT на более доступном оборудовании – обычном смартфоне.

Цель данной работы – подобрать оптимальный метод определения положения головы, а также оптимальный алгоритм фильтрации данных положения головы. Также в работе будут приведены характеристики используемой камеры для записи эксперимента vHIT с целью определения положения головы пациента. В настоящей статье мы введем метрики, оценивающие точность определения положения головы пациента, а также введем пороговое значение, ниже которого точность определения является неудовлетворительной. Это пороговое значение важно в контексте предъявления требований к алгоритмам определения головы для их использования в vHIT-тесте.

Структура статьи такова. В разд. 2 сравниваются текущие актуальные решения для записи и оценки vНИТ-теста. В разд. 3 подробно описаны условия постановки теста vНИТ в лабораторных условиях для набора данных с целью извлечения из записанных видео информации о положении головы в каждый момент времени. В разд. 4 проведено сравнение результатов работы двух моделей, обученных на данных IMU, и волоконного сенсора. В разд. 5 приведена сводная таблица результатов нашего сравнения, а также сделан вывод о том, какой метод фильтрации и для какой модели машинного обучения является наиболее точным.

2. СРАВНЕНИЕ ТЕКУЩИХ РЕШЕНИЙ

Наиболее популярные системы, позволяющие записывать и оценивать vНИТ-тест, – это Interacoustics, Otometrics и Synapsys [1]. Авторы статьи [1] поставили цель сравнить эти три устройства, чтобы понять, есть ли согласованность между результатами систем на предмет постановки диагноза билатеральная вестибулопатия (bilateral vestibulopathy, BV). Все эксперименты vНИТ с тремя системами проводил один и тот же специалист. Расстояние от спинки стула пациента до точки фиксации составило два метра. Комната была хорошо освещена, чтобы зрачок у пациента имел малые размеры. В качестве точки фиксации использовался лазер с длиной волны 532 нм. Точка фиксации была отрегулирована на уровне глаз каждого пациента. Системы Interacoustics и Otometrics сделаны в виде очков с ремешком, на одном из глаз (правом) установлена камера для фиксации движения зрачков. Система Synapsys в виде камеры расположена перед пациентом, она фиксирует движения обоих глаз (рис. 1).

Рис. 1. Схемы установки система Interacoustics, Otometrics и Synapsys.

Главным параметром, который использовался для сравнения систем, был выбран коэффициент вестибулоокулярного рефлекса (VOR gain). Система Interacoustics вычисляет мгновенный коэффициент VOR в малом окне по времени (порядка 60 мс). В этом временном окне вычисляется скорость движения глаз и головы. Система Otometrics вычисляет коэффициент усиления VOR как отношение площади под кривой скорости глаз и скорости головы. Временной диапазон берется от 60 мс до пикового ускорения головы до возвращения головы в состояние покоя. Вся информация о Synapsys – это то, что она вычисляет коэффициент усиления VOR за период от 40 мс до пикового ускорения головы до 80 мс после пикового ускорения. В случае скрытой саккады 80-миллисекундное окно сокращается и останавливается во время начала скрытой саккады. Частота кадров у системы Synapsys составляет 100 Гц. Для увеличения точности детектирования была

применена линейная интерполяция, увеличивающая частоту кадров до 250 Гц. Для сравнения, частота кадров у систем Interacoustics и Otometrics составляет 220 и 245 Гц соответственно.

У 28 пациентов, протестированных со всеми тремя системами vHIT, процент соответствия результатов vHIT составил 79% (68% BV, 11% без BV), а у 21% соответствия не было. Средние коэффициенты VOR, полученные у этих 28 пациентов, составили 0.36, 0.36 и 0.09 для Interacoustics, Otometrics и Synapsys соответственно. Во всех результатах системы Interacoustics и Otometrics были близки по измеренным значениям друг к другу, а результаты Synapsys сильно отличались.

Поскольку в нашем случае вновь разрабатываемая система будет ближе по конструкции к Synapsys, нужно, опираясь на аппаратные характеристики данного устройства, предъявить требования к разрабатываемой системе. Итак, известно, что частота кадров Synapsys (100 Гц) не обеспечивает должного уровня согласованности с двумя остальными системами, а значит, показывает точность определения коэффициента VOR и скорости вращения головы ниже, чем у аналогов. Очевидно, что чем выше частота кадров, тем выше точность распознавания положения головы и зрачков. Вторым важным аспектом является разрешение камеры, а точнее, какое количество пикселей приходится на зрачок пациента. В случае с Interacoustics и Otometrics в поле зрения камеры находится лишь один глаз, что позволяет иметь очень высокую точность определения положения зрачка. На сайте производителя Synapsys видно, что разрешение камеры составляет 1456×1080 пикселей. Полагаем, что передняя поверхность линзы объектива камеры расположена на расстоянии примерно 1 м. В табл.1 приведены характеристики камер коммерчески доступных смартфонов.

Таблица 1. Сравнение характеристик передних камер смартфонов

Название смартфона	Частота кадров передней камеры, Гц	Разрешение передней камеры, пиксель
iPhone 11–16	120	1920×1080
Samsung Galaxy s24/s24+/s24 ultra	60	3840×2160
Huawei Pura 70 ultra	120	1920×1080
Xiaomi	60	3840×2160

Исходя из данных табл. 1, важно понять, какой фактор наиболее критично влияет на качество распознавания движения головы и зрачков – частота кадров или разрешение камеры. Мотивацией того, что для рассмотрения была выбрана передняя камера, является

то, что в этом случае пациент или врач сможет оперативно видеть рекомендательные сообщения на экране смартфона для корректировки vHIT-теста. Также на экране смартфона можно расположить точку, на которой пациент должен держать взгляд.

3. ОБЩЕЕ ОПИСАНИЕ ПОСТАНОВКИ ЭКСПЕРИМЕНТА ДЛЯ ЗАПИСИ ДАННЫХ vHIT С ПОМОЩЬЮ КАМЕРЫ СМАРТФОНА

Определим основные факторы, которых необходимо придерживаться в ходе записи данных vHIT-теста. Очевидно, что они могут корректироваться в зависимости от результата работы алгоритмов детектирования движения зрачков и головы пациента, однако необходимо на данном этапе выработать первую версию сценария эксперимента. Общий план эксперимента представлен на рис. 2.

Рис. 2. Общая схема постановки эксперимента по записи данных vHIT-теста на смартфон.

Исходя из рис. 2, в общем случае мы имеем всего два параметра расстояния. Основываясь на опыте Synapsys, назначим расстояние $L=1$ м, а расстояние H будем подбирать для каждого пациента таким образом, чтобы оптическая ось объектива камеры совпадала с оптической осью глаза пациента.

При первом сборе данных мы будем использовать фронтальную камеру iPhone 12 Pro – на этих данных будет обучаться оптимально подобранная модель машинного обучения. Далее необходимо подобрать опорные датчики, которые будут предоставлять истинную информацию о положении зрачков и головы пациента в каждый момент времени. В случае с положением зрачков пациенту довольно сложно использовать стороннее устройство, за исключением подхода, описанного в работе [2]. Авторы сравнивали такие подходы к распознаванию движения зрачков пациента, как видеоаналитика и склерические кольца. Авторы утверждают, что точность видеоаналитики хорошо согласуется с точностью склерических колец за исключением случаев с движением головы относительно камеры. За неимением склерических колец, а также по соображениям сложности использования данного решения для записи большого объема данных, было решено использовать видеоаналитику с самым точным методом определения положения зрачков. Мы можем терять в скорости обработки видеоданных, однако при создании опорных данных нам важна точность, и ради нее мы готовы пожертвовать временем обработки. В случае с подбором опорных датчиков для предоставления априорной информации о положении головы пациента есть два варианта: датчики IMU и волоконный сенсор (рефлектометр). Мы будем собирать часть данных с IMU-сенсорами в качестве опорных, а часть – с рефлектометром.

4. ВЫЧИСЛЕНИЕ ПОЛОЖЕНИЯ ГОЛОВЫ ПАЦИЕНТА

Для начала необходимо определиться, в каком виде будет предоставлена информация о положении головы. Поскольку необходимо сопоставлять движение зрачков и движение головы на одном графике, логично представлять данные о положении и перемещении головы в виде текущих углов рыскания, тангажа и крена. Источником опорных данных может служить множество датчиков. Самое распространенное устройство для записи таких данных – сенсор IMU (inertial measurement unit), предоставляющий данные о линейном ускорении, а также об угловой скорости по трем осям. Также можно использовать данные, например с оптического дальномера, он будет предоставлять информацию только о текущем расстоянии до стены, но зная расстояние L при нулевом положении головы (рис. 3) и текущее расстояние, легко можно рассчитать текущий угол рыскания. Плюсом такого метода могут стать высокая точность и простота установки измерителя на голове пациента, а минусом – быстродействие и невозможность разделения данных на три оси.

Рис. 3. Схема расчета угла рыскания α

4.1. Использование OFDR в качестве источника опорной информации о положении головы

В распоряжении команды был волоконно-оптический сенсор OFDR (optical frequency domain reflectometer). В качестве начального варианта для измерения угловых положений головы он был приспособлен к работе в качестве оптического дальномера путем установки коллиматора на выходе оптической линии. На голову пациента был установлен коллиматор. На стене были предварительно отмечены пять точек (рис. 3). Расстояния между всеми соседними точками одинаковые (112 см). Расстояние L_0 от коллиматора до нулевой точки составило 186 см. Было записано пять рефлектограмм для каждой точки (рис. 4).

Рис. 4. Рефлектограмма, записанная при положении головы P_2

Как видно на рис. 4, для каждой рефлектограммы нам важно вычислить расстояние между двумя пиками: первый характеризует конец оптической системы, а второй – пик отражения от стены. Итак, результат простейших вычислений сведен в табл. 2.

Таблица 2. Результаты эксперимента по применению OFDR в качестве дальномера

Положение	Измеренное расстояние, см	Измеренный угол рыскания α
M_2	294.7	50.6°
M_1	220.5	32.2°
0	186	0
P_1	215.6	29.9°
P_2	289.6	49.8°

На точность в данном случае влияет то, насколько горизонтально пациент может перемещать голову между двумя соседними точками. Как следует из результата в табл. 2, небольшое различие углов рыскания как раз может быть связано с тем, что, например, при перемещении между точками P_1 и P_2 присутствовал небольшой вертикальный поворот.

4.2. Использование IMU-модуля в качестве источников опорной информации положения головы

Второй эксперимент был произведен в точном соответствии с описанием в разд. 3. Данные с камеры смартфона iPhone 12, а также с IMU-сенсоров записывались одновременно. В качестве устройства IMU был использован смартфон iPhone 5, установленный на затылке пациента таким образом, чтобы он не попадал в кадр. Данные забираются одновременно со встроенных трехосевого акселерометра и гироскопа. Для синхронизации данных с камеры и с IMU были использованы звуковые сигналы, издаваемые смартфоном с IMU, обозначающие начало и конец эксперимента. Частота кадров с передней камеры составила 120 Гц, а частота дискретизации данных с IMU – 98 Гц, поэтому была проведена интерполяция данных с IMU-сенсоров. Результат эксперимента показан на рис. 5.

Рис. 5. Результаты синхронизации данных с камеры смартфона и IMU сенсоров.

Сенсоры IMU дают более полную картинку положения головы, а также имеют относительно высокую частоту дискретизации при записи опорных данных в сравнении с оптическим дальномером.

4.3. Применение алгоритмов определения положения головы пациента

В интерпретации vHIT-теста неотъемлемой частью является детектирование положения зрачков пациента, причем положения зрачков должны выдаваться не относительно положения камеры, а относительно положения головы пациента. В связи с этим точность определения положения головы пациента должна быть максимальной.

Для того чтобы получить полезную информацию о положении головы пациента, нужно обработать видеопоток, записанный на камеру смартфона. На начальном этапе необходимо проверить, насколько готовые алгоритмы могут справляться с задачей определения положения головы. Чтобы провести сравнение алгоритмов на предмет их состоятельности и точности, нужно определиться с методологией оценки. Как было описано выше, у нас есть опорные сигналы линейного ускорения и угловой скорости по трем осям. Можно выделить следующие основные этапы сравнения опорных данных с IMU и данных на выходе готовых алгоритмов обработки видеопотока.

1. Определение размерности данных на выходе готовых алгоритмов.
2. Приведение опорных данных и данных после готовых алгоритмов к единой размерности.
3. Фильтрация данных после готовых алгоритмов.
4. Выбор метрики сравнения опорных и полученных данных.

Было принято решение провести сравнение трех готовых моделей машинного обучения (МО): DHPE (dense head pose estimation), HPE (head pose estimation), LHPE (lightweight head pose estimation) [3–5]. В их основе лежат сверточные нейронные сети, структуры которых различаются внутри каждой из этих моделей.

Для фильтрации данных проведено сравнение трех популярных методов. Первый метод – это эмпирическая модовая декомпозиция (empirical mode decomposition, EMD). Это популярный инструмент анализа данных, который позволяет учитывать локальные особенности сигнала, а также наличие помех различного рода [6–8]. Главным образом, EMD способен анализировать нестационарные и нелинейные сигналы, что релевантно нашему применению. Второй, не менее популярный метод – это вариационная модовая декомпозиция (variational mode decomposition, VMD). Этот метод также показал себя эффективным по части разделения исходного сигнала на моды, выделения полезного сигнала и шума [9, 10]. В отличие от EMD метод VMD более устойчив к фоновому шуму [11], что может увеличить его эффективность в нашем случае, так как влияние фонового шума на выходе DHPE, HPE и LHPE на данном этапе непонятно. Также для данных задач фильтрации широко распространено дискретное вейвлет-преобразование [12], результаты применения которого мы также приведем. Наконец, четвертый метод – это фильтрация Баттервортта. Этот фильтр тоже достаточно распространен в силу того, что его амплитудно-частотная характеристика (АЧХ) не имеет пиков в полосе пропускания, а представляет собой гладкую кривую [13]. В нашем случае этот фильтр выступает как некий базовый инструмент цифровой обработки сигналов.

4.3.1. Определение размерности данных на выходе моделей детектирования положения головы. После применения моделей DHPE, HPE и LHPE к записанному видеоролику на выходе из данных моделей мы получили размерности выходных данных (табл. 3). Сравним их с опорными данными (табл. 4).

Таблица 3. Структура и размерность выходных данных из моделей детектирования положения головы

Модель	Выходные данные	Размерность выходных данных
DHPE	Углы Эйлера	Градусы
HPE	Углы Эйлера	Градусы
	Вектор смещения (3 оси, определяет соответствие между мировой системой координат и системой координат камеры)	Пиксели
LHPE	Углы Эйлера	Градусы

Таблица 4. Структура и размерность опорных данных IMU

IMU-сенсор	Выходные данные	Размерность выходных данных
Акселерометр	Угловые ускорения (3 оси)	рад/с ²
Гироскоп	Угловые скорости (3 оси)	рад/с

Нам полезны угловые скорости в качестве выходных данных МО. Как видно из сравнения, для приведения размерности данных, полученных с помощью моделей DHPE, HPE и LHPE, к размерности опорных данных, нужно провести нормировку данных и продифференцировать полученные данные. Приведем полный алгоритм преобразования полученных данных после модели МО (рис. 6).

Рис. 6. Преобразование данных на выходе МО. Подходит для DHPE, HPE, LHPE.

Выходные данные моделей МО будут отличаться друг от друга и, соответственно, от опорных данных, на это также будет влиять цифровой частотный фильтр (рис. 6). Поэтому далее будет приведено сравнением разных методов фильтрации (нам надо будет продемонстрировать самый эффективный метод). В качестве метрики сравнения выберем MSE (mean squared error):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2, \quad (1)$$

где n – количество точек в сигнале, y_i – опорный сигнал IMU, \tilde{y}_i – преобразованный сигнал. Это распространенная метрика во многих областях, она выступает как функция ошибок во многих моделях машинного обучения [14, 15].

Стоит отметить, что сигналы с IMU-сенсоров нормируются до диапазона значений от 0 до 1 (рис. 6) (X_{scaled}), чтобы все данные находились в одном масштабе. Нам принципиально важно понять, совпадает ли форма опорных сигналов и сигналов после моделей МО, а для этого абсолютными значениями сигналов можно пренебречь.

4.3.2. Определение критических значений MSE для эксперимента vHIT. Для более информативного представления результатов необходимо ввести пороговый критерий значений MSE, выше которого мы будем считать, что результат эксперимента vHIT имеет неудовлетворительную точность. Для этого мы используем опорный сигнал с IMU-гироскопа со внесенными в него искажениями. Будем использовать следующую методику внесения искажений: определим диапазон коэффициентов, на которые мы будем умножать каждое значение опорного сигнала IMU. Диапазон коэффициентов искажения напрямую зависит от динамического диапазона сигнала. Для каждой точки опорного сигнала значение коэффициента искажения будем выбирать из установленного диапазона случайно. Формула для расчета коэффициента α_i для i -й точки сигнала такова:

$$\alpha_i = s_i (d_i + rn_i), \quad (2)$$

где s_i – значение сигнала в i -й точке, d_i – динамический диапазон сигнала, rn_i – значение случайного числа, выбранного в установленном диапазоне коэффициентов искажения.

Для того чтобы оценивать эффективность того или иного метода фильтрации, мы будем использовать выражение (1), однако учтем тот факт, что значение размаха сигналов в абсолютных величинах по оси x будет намного больше, чем таковые по осям y и z . Исходя из реалий проведения теста vHIT, понятно, что основной вклад вращения головы будет только по одной оси – примем за эту ось за x (рис. 7).

Рис. 7. Положение осей x, y, z в системе координат головы пациента

Две другие оси имеют сравнительно малые абсолютные размахи сигнала (табл. 5). Поэтому сравнивать методы фильтрации сигналов на выходе моделей МО мы будем только по одной оси – x . Однако из табл. 5 также видно, что вклады ошибок MSE по оси x для трех методов также будут различаться в силу того, что абсолютные размахи сигналов у трех методов немного различаются. Поэтому мы будем использовать значения размахов сигналов как весовой коэффициент, на который мы будем умножать MSE, тем самым, мы получим взвешенные MSE, которые уже можно сравнивать между собой.

Таблица 5. Значения абсолютных размахов сигналов угловой скорости IMU при повороте головы по трем осям, используемые в качестве нормировочных коэффициентов

Модель МО, ось	Абсолютный размах сигнала	Среднее значение размаха
DHPE, x	95.52°	85.91°
HPE, x	70.6°	
LHPE, x	91.61°	
DHPE, y	23.9°	21°
HPE, y	16.78°	
LHPE, z	22.32	
DHPE, z	15.18°	14.45°
HPE, z	12.82°	
LHPE, z	15.36°	

После искажения опорного сигнала мы считаем MSE и полученную MSE умножаем на среднее значение размаха сигнала по оси (табл. 5). На искажения сигнала могут влиять как сама модель МО, так и условия записи видео: ракурс, освещение, положение головы в кадре. Выберем стандартное отклонение для нормального распределения 0.95 (рис. 8). В качестве тестового сигнала для искажения был взят сигнал с гироскопа IMU по оси x (рис. 9).

Рис. 8. Гистограмма распределения коэффициентов искажения

Рис. 9. Кривые опорного и искаженного сигналов гироскопа IMU по оси x

После расчета взвешенный MSE составил 0.7. Это значение примем за пороговое в индикации удовлетворительности проведения теста vHIT.

4.3.3. Использование вариационной модовой декомпозиции в обработке выходных данных моделей детектирования положения головы. Рассмотрим алгоритм VMD (variational mode decomposition), который является адаптивным методом разложения сигнала на конечное число дискретных мод, каждая из которых имеет ограниченный частотный диапазон. Метод формулируется как задача оптимизации, где одновременно извлекаются моды и их соответствующие центральные частоты. В отличие от эмпирической модной декомпозиции (EMD), метод VMD обладает устойчивостью к шуму и менее подвержен эффекту наложения мод (mode mixing). Он использует вариационный подход с применением преобразования Вейерштрасса и метода множителей Лагранжа. Метод VMD широко применяется в анализе нестационарных сигналов, таких как биомедицинские данные, вибросигналы и другие.

Поскольку алгоритм VMD имеет множество параметров, обратимся к литературе, где для похожей задачи фильтрации указаны оптимальные настройки для VMD [16]. В данной работе описывается задача анализа акустических сигналов с целью обнаружения дефектов в дуговых магнитах. Хотя в нашей работе анализу подвергается не акустический сигнал, в нем также содержится множество мод, которые характеризуют состав сигнала. Метод VMD применяется для предварительной фильтрации и выделения информативных компонент. Оптимальные параметры алгоритма VMD из работы [16] отображены в табл. 6.

Таблица 6. Установленные параметры VMD-алгоритма

Параметр – описание	Значение
alpha – параметр ограничения точности данных	1636
tau – шаг двойного подъема	0.0
K – количество восстанавливаемых мод	3
DC – учет нулевой частоты в сигнале	0
tol – допуск критерия сходимости	10^{-6}

Метод VMD используется в качестве цифрового частотного фильтра, каждая из полученных мод обрабатывается согласно предложенному алгоритму (рис. 6), а затем ведется расчет взвешенного MSE, после чего выбирается наименьшая MSE для каждой модели МО по оси x . Приведем графики зависимости MSE по оси x от номера IMF (рис. 10).

Рис. 10. График зависимости взвешенных MSE от номера моды, полученной алгоритмом VMD, примененным к выходным сигналам из моделей МО DHPE, LHPE, HPE по оси x

Наилучшие значения взвешенной MSE = 0.91 для IMF №1 (DHPE). Результаты выделения оптимальных IMF из сигналов приведены на рис. 11.

Рис. 11. Сигналы DHPE, LHPE, HPE (после применения алгоритма VMD, выбраны оптимальные IMF) и гироскопа IMU по оси x

Полученное значение взвешенного MSE по оси x выше, чем пороговое значение. Следовательно, точности VMD недостаточно для обработки эксперимента vHIT при данных параметрах алгоритма VMD. Для получения более высокоточного результата и, как следствие, меньшего значения MSE, требуется проведение отдельной исследовательской работы по оптимизации параметров алгоритма VMD.

4.3.4. Использование эмпирической модовой декомпозиции в обработке выходных данных моделей детектирования положения головы. Метод EMD представляет собой

итерационное разложение входного сигнала на эмпирические моды (IMF, intrinsic mode function). В основе алгоритма лежит построение гладких верхней и нижней огибающих, вычисление средних значений между верхней и нижней огибающими, а затем вычитание из входного сигнала этих средних значений. Процесс является итерационным, пока не будет достигнут расчет всего количества n IMF, заданного пользователем. Пусть у нас есть входной сигнал $X_{in}(t)$. Разложение этого сигнала методом EMD – это вычисление эмпирических мод c_j и остатков $r_j = r_{j-1} - c_j$, где $j = 1, \dots, n$, $r_0 = X_{in}(t)$. Таким образом, формула разложения сигнала методом EMD выглядит следующим образом:

$$X_{in}(t) = \sum_{j=1}^n c_j + r_n. \quad (3)$$

Получение остатков r_n происходит до тех пор, пока вновь выделенный остаток не станет монотонной функцией, из которой эмпирическую моду выделить уже не представляется возможным. В реальных применениях, в том числе и в цифровой обработке сигнала, количество IMF может быть ограничено пользователем.

Применим EMD для выделения полезного сигнала из сигналов, сформированных на выходе модели МО. Напомним, что поскольку частота кадров камеры, на которую было записано видео vHIT, составляет 120 Гц, сигнал положения головы по трем осям, полученный с помощью моделей МО, также имеет частоту дискретизации 120 Гц. В разд. 4.2 было описано, что сигналы с IMU были интерполированы до частоты 120 Гц. В EMD есть важный параметр, который определяет, при достижении какого значения амплитуды остатка процесс фильтрации прекращается. По умолчанию значение данного параметра равно 10^{-8} . В реалиях наших амплитуд, в том числе нормированных (от 0 до 1), это значение является релевантным. При таком значении максимальное количество IMF при применении EMD к выходному сигналу любой модели МО $n = 8$. Исходя из этого, выберем максимальное количество мод, равное 8.

Последовательность исследований метода EMD показана на рис. 6, где EMD выступает в роли цифрового частотного фильтра. После EMD мы имеем n функций IMF, поэтому для каждой из них проводятся преобразования, указанные после цифрового частотного фильтра. В результате мы получим n сигналов X_{out} , для каждого из которых далее необходимо посчитать взвешенное значение MSE для каждого из n сигналов.

Рис. 12. Графики зависимости взвешенных MSE от номера IMF, полученной алгоритмом EMD, примененным к выходным сигналам из моделей МО DHPE, LHPE, HPE по оси x

Результат анализа алгоритма EMD отображен на рис. 12. Минимальные значения взвешенной MSE приходятся на IMF №2, причем сразу для трех моделей МО. Наилучшие значения взвешенной MSE получены у модели LHPE, MSE = 0.89 для IMF №2.

На рис. 13 видно, что LHPE ближе всех к опорному сигналу, особенно по оси x , однако отличие не является большим.

Рис. 13. Сигналы DHPE, LHPE, HPE (после применения алгоритма EMD, оптимальная IMF №2) и гироскопа IMU по оси x

Полученные значения MSE по оси x меньше, чем в случае с VMD, однако метод EMD все же не удовлетворяет пороговому значению MSE, определяющему качество обработки эксперимента vHIT.

4.3.5. Применение вейвлет-преобразования для обработки выходных данных моделей детектирования положения головы. Фильтрация сигнала с помощью вейвлет-функции начинается с выбора подходящего типа функции. Выбор должен быть сделан с учетом конкретных характеристик шума и сигнала, а также желаемого результата шумоподавления; однако рекомендуется протестировать несколько вейвлет-функций, поскольку они могут демонстрировать различное поведение в зависимости от конкретного случая.

Когда к сигналу применяется выбранное вейвлет-преобразование, оно работает как последовательность фильтров верхних и нижних частот, что приводит к разложению сигнала на несколько поддиапазонов, каждый из которых представляет различные частотные диапазоны. Эта процедура позволяет получить вейвлет-коэффициенты, классифицируемые либо как аппроксимирующие (cA), отвечающие за захват низкочастотной информации, либо как детализирующие (cD), отвечающие за захват высокочастотной информации (рис. 14) [12].

Рис. 14. Схема обработки сигнала с помощью вейвлет-преобразования

Исходя из работы, в которой описывается применение вейвлет-преобразования для схожей задачи, а именно фильтрации сигналов с сенсоров IMU, оптимальной функцией является вейвлет-функция Добеши с уровнем декомпозиции 6 [17]. В приведенной статье не указан порядок этой функции, поэтому мы применяли для разложения несколько различных вейвлет-функций, а именно db1, db2, db4, db8. Также в работе [17] приведена формула для расчета порогового значения для исключения получившихся коэффициентов после вейвлет-преобразования. Мы приняли решение ввести следующий дополнительный коэффициент для регулировки порога:

$$\lambda = \sqrt{2\lg(n)\sigma} \cdot \alpha, \quad (4)$$

где n – количество точек в сигнале, σ – значение стандартного отклонения сигнала, α – коэффициент регулировки порога. Значения коэффициента регулировки порога выбраны следующими: 0.1, 0.2, 0.4, 0.8, 1.

После применения вейвлет-преобразования с приведенными ранее параметрами мы получили результаты расчета взвешенных MSE, которые свели в табл. 7.

Таблица 7. Взвешенные MSE после применения вейвлет-преобразования по оси x

db1			
A	DHPE	LHPE	HPE
0.1	1.844	1.348	1.747
0.2	1.905	1.401	1.811
0.4	1.975	1.454	1.886
0.8	2.011	1.481	1.921
1	2.011	1.481	1.921
db2			
A	DHPE	LHPE	HPE
0.1	0.1	1.166	0.811
0.2	0.2	1.363	0.982
0.4	0.4	1.708	1.263
0.8	0.8	1.907	1.398
1	1	1.907	1.399
db4			
A	DHPE	LHPE	HPE
0.01	0.599	0.095	0.16
0.05	0.388	0.214	0.28
0.1	0.531	0.364	0.464
0.2	0.912	0.85	0.936
0.4	1.989	1.708	2.015
0.8	2.611	1.906	2.476
1	2.631	1.906	2.499
db8			
A	DHPE	LHPE	HPE
0.1	0.775	0.424	0.535
0.2	1.601	1.184	1.207
0.4	2.304	1.994	2.351
0.8	3.413	2.588	3.248
1	3.421	2.588	3.322

Для вейвлет-функции db4 мы взяли дополнительные точки коэффициента α (0.05 и 0.01), и в точке 0.05 мы нашли локальный минимум взвешенной MSE. Для каждой модели МО выбрано минимальное значение взвешенной MSE, означающее наилучший результат работы вейвлет-преобразования.

Итак, с учетом табл. 7 по оси x наилучший результат применения вейвлет-преобразования (наименьшая MSE) наблюдается для модели LHPE (0.095). Приведем вид сигнала для наилучших конфигураций фильтрации (рис. 15).

Рис. 15. Сигналы LHPE и HPE (после применения алгоритма вейвлет-преобразования) с наименьшей взвешенной MSE и сигналы гироскопа IMU по оси x

Значения взвешенной MSE меньше, чем установленный порог MSE, что говорит о том, что метод вейвлет-преобразования справился с задачей фильтрации.

4.3.6. Использование фильтра Баттерворт для обработки выходных данных моделей детектирования положения головы. Основными параметрами фильтра Баттервортя являются его тип, частота среза и его порядок. Все они влияют на форму АЧХ этого фильтра. Тип фильтра задается, исходя из того, в какой частотной области мы хотим пропустить сигнал, а в какой отфильтровать. В зависимости от этого бывают фильтры низких частот, высоких частот и полосовые фильтры. Частоты среза в данном случае показывают частотные границы фильтра, на которых происходит снижение кривой АЧХ на 3 дБ. Порядок фильтра определяет крутизну спада АЧХ на частотах среза. Зададим начальные параметры фильтра Баттервортя, которым отфильтруем сигнал, сформированный моделями МО, а затем возьмем производную согласно рис. 6. Начальные параметры фильтра сведены в табл. 8.

Таблица 8. Начальные параметры фильтра Баттервортя

Параметр	Значение/Комментарий
Тип фильтра	Нижние частоты
Частота среза, Гц	10
Порядок фильтра	3

Очевидно, что мы хотим отфильтровать высокочастотные шумы от работы моделей МО, поэтому мы выбрали именно такой тип фильтра. Исходя из записанного vHIT, частота вращения головы пациентов чуть больше 1 Гц. Однако с учетом неравномерных угловых ускорений головы, а также резких остановок (отрицательных ускорений) при крайних положениях головы мы решили взять частоту 10 Гц, чтобы не отфильтровать эти полезные для анализа высокочастотные составляющие.

Приведем на одном графике сигналы для DHPE, HPE, LHPE и гироскопа IMU (рис. 16).

Рис. 16. Сигналы DHPE, LHPE, HPE (после фильтрации Баттервортя) и гироскопа IMU по оси x

Для выбранного значения частоты среза мы посчитали все взвешенные MSE для каждой модели МО (рис. 17).

Рис. 17. Значения взвешенных ошибок MSE по оси x

Возможно, параметры фильтра (табл. 8) неоптимальны. Поэтому было изучено изменение нормированных ошибок MSE при разных частотах среза (рис. 18).

Рис. 18. Графики зависимости взвешенных MSE от изменения частоты среза фильтра Баттервортса по оси x

Значения оптимальных частот среза фильтра по оси x для DHPE, LHPE, HPE равны соответственно 7, 8, 7 Гц. На рис. 19 представлен график отфильтрованных сигналов с оптимальными частотами среза фильтра Баттервортса.

Рис. 19. Сигналы DHPE, LHPE, HPE (после фильтрации Баттервортса с оптимальными частотами среза) и гироскопа IMU по оси x

Можно сделать вывод о том, что при оптимальной фильтрации наилучшее совпадение с формой опорного сигнала гироскопа по всем трем осям наблюдается у HPE и LHPE (рис. 17). Наилучшие значения взвешенной MSE по оси x составляет 0.13.

Фильтрация Баттервортса является одной из самых точных из всех перечисленных методов: значение MSE по оси x меньше порогового, а значит, фильтр Баттервортса в данных условиях является релевантным для фильтрации данных на выходе моделей МО.

5. ОБСУЖДЕНИЕ

Приведем сводную таблицу взвешенных MSE для каждого метода фильтрации и для каждой модели МО (табл. 9).

Таблица 9. Итоговая таблица всех взвешенных MSE

Модель МО	Метод фильтрации	Значение взвешенной MSE
DHPE	VMD	0.91
LHPE		0.93
HPE		1.08
DHPE	EMD	1.48
LHPE		0.89
HPE		1.11
DHPE	Вейвлет-преобразование	0.388
LHPE		0.095
HPE		0.16
DHPE	Фильтр Баттервортса	0.33
LHPE		0.13
HPE		0.13

Отметим, что методы вейвлет-преобразования и фильтрации Баттервортса имеют наименьшие взвешенные MSE, а значит, результат фильтрации лучше совпадает с опорным сигналом сенсора IMU. Также следует добавить, что выбранное значение $\text{fps}=120$ Гц и разрешение 1920×1080 являются достаточными для того, чтобы получить сигнал поворота головы после всех стадий цифровой обработки сигнала на выходе моделей МО, который по форме хорошо совпадает с опорным сигналом.

Результаты, полученные в этой статье, показывают потенциал для популяризации проведения vНИТ-тестов, в частности, из-за того, что продемонстрирована высокая степень совпадения кривых углового ускорения, полученных от опорного датчика IMU и после представленного алгоритма обработки сигналов от моделей МО. Эти кривые непосредственно используются как часть основной информации для интерпретации результатов теста vНИТ. Еще одна часть основной информации – это кривые перемещения зрачков в системе координат, совпадающей с системой координат головы пациента. Задача вычисления таких кривых только с помощью камеры смартфона и цифровой обработки – большой вызов, который может накладывать более жесткие требования к условиям эксперимента, а также к характеристикам камеры. Возможно, в условия эксперимента необходимо будет ввести дополнительную активную подсветку для того, чтобы зрачки были легче детектируемы и отслеживаемы с помощью различных алгоритмов. Более того, возможно, потребуется увеличение fps-камеры как на фундаментальном уровне (замена устройства), так и на алгоритмическом уровне (интерполяция кадров).

6. ЗАКЛЮЧЕНИЕ

Подводя итоги, сформулируем основные результаты работы. vНИТ является одним из современных методов диагностики больных, обратившихся с жалобами на головокружения. Для успешной диагностики необходимо правильно определять положение как головы, так и зрачков пациента при резком повороте головы с малой амплитудой. Коммерческие системы продемонстрировали свою пригодность для решения этой задачи, но они малодоступны ввиду высокой стоимости. Авторы ставят перед собой задачу сконструировать систему для проведения vНИТ-теста с помощью подручных средств, а именно – камеры и датчиков смартфона.

Первым шагом на пути построения системы является ответ на вопрос, возможно ли определение положения головы по видеосъемке с нужной точностью. Этому и посвящена данная работа.

Авторы разработали критерий успешности определения положения головы. Для этого брался идеальный сигнал (зависимость координаты головы от времени), и в него вносились искусственные помехи. По мере роста помех возрастало среднеквадратичное отклонение MSE (нормированного) возмущенного сигнала от исходного. Таким образом было определено пороговое значение MSE, оно составило 0.7.

Затем были протестированы различные методы определения положения головы и последующей цифровой фильтрации сигнала, полученного из видеопотока, снятого камерой смартфона iPhone 12 Pro. При этом одним из факторов, влияющих на качество полученного сигнала, является частота дискретизации кадров.

Наилучшим с точки зрения минимизации MSE методом определения положения головы оказался метод LHPE. Также приемлемые результаты показали и другие основанные на сверточных нейронных сетях методы – HPE и DHPE.

Наилучшими алгоритмами фильтрации сигнала оказались методы вейвлет-преобразования и фильтрации Баттервортса. В то же время, методы VMD и EMD дают значения MSE больше порогового, т. е. не годятся для решения поставленной задачи.

Использование оптимального сочетания LHPE и вейвлет-преобразование позволяет достичь значения MSE менее 0.1, что почти на порядок меньше порогового. Близкие, хоть и чуть большие, значения MSE дают и другие комбинации методов, представленные в табл. 9.

Таким образом, задачу определения с нужной точностью положения головы при vHIT-тесте с помощью камеры смартфона можно считать решенной. Следующим шагом должно стать решение задачи определения положения зрачка с помощью того же оборудования.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках Государственного задания № 124020600009-2.

СПИСОК ЛИТЕРАТУРЫ

1. *van Dooren T.S., Starkov D., Lucieer F.M.P. et al.* // J Neurol. 2020. V. 267. P. 256. <https://link.springer.com/article/10.1007/s00415-020-10060-w>
2. *Geest J.N., Frens Maarten.* // J. Neurosci. Methods. 2002. V. 114. P. 185. [http://doi.org/10.1016/S0165-0270\(01\)00527-1](http://doi.org/10.1016/S0165-0270(01)00527-1)
3. <https://github.com/1996scarlet/Dense-Head-Pose-Estimation>

4. <https://github.com/yakhyo/head-pose-estimation>
5. <https://github.com/Shaw-git/Lightweight-Head-Pose-Estimation>
6. *Belokrylov M.E., Kambur D.A., Konstantinov Y.A., Claude D., Barkov F.L.* // Sensors. 2024. V. 24. P. 1253. <http://dx.doi.org/10.3390/s24041253>
7. *Qin Z., Chen H., Chang J.* // Sensors. 2017. V. 17. P. 1870. <http://doi.org/10.3390/s17081870>.
8. *Barbosh Mohamed, Singh Premjeet, Sadhu Ayan.* // Smart Mater. Struct. 2020. V. 29. № 9. P. 093001. [10.1088/1361-665X/aba539](https://doi.org/10.1088/1361-665X/aba539)
9. *Sinam Ajitkumar Singh, Sinam Ashinikumar Singh, Ningthoujam Dinita Devi, Swanirbhar Majumder.* // Electronic Devices, Circuits, and Systems for Biomedical Applications. / Ed. by Suman Lata Tripathi, Valentina E. Balas, S.K. Mohapatra, Kolla Bhanu Prakash, Janmenjoy Nayak. Academic Press, 2021. Ch 7. P. 135. <https://doi.org/10.1016/B978-0-323-85172-5.00016-2>
10. *Liu Jun-Zhou, Xue Ya-juan, Shi Lei, Han Lei.* // Front. Earth Sci. 2022. V. 10. P. 1. <https://doi.org/10.3389/feart.2022.917747>
11. *Uzunoğlu Cengiz. A.* // Electrica. 2018. V. 18. P. 72. <http://dx.doi.org/10.5152/ijeee.2018.1811>
12. *Hasan A.M., Samsudin K., Ramli A.R., Azmir R.S.* // J. Appl. Sci. 2010. V. 10. P. 2217. <https://doi.org/10.3923/jas.2010.2217.2230> .
13. *Ellis G.* // Control System Design Guide. / Ed. by G. Ellis. Butterworth-Heinemann, 2012. Ch. 9. P. 165.
14. *Martin M., Dognin C., Shu Y., Blot M.* // ArXiv:1908.02718. <https://doi.org/10.48550/arXiv.1908.02718> .
15. *Botchkarev A.* Evaluating Performance of Regression Machine Learning Models Using Multiple Error Metrics in Azure Machine Learning Studio. 2018. <https://doi.org/10.2139/ssrn.3177507>
16. *Huang Q. Liu X., Li Q., Zhou Y., Yang T., Ran M.* // AIP Adv. 2021. V. 11. P. 065216. <https://doi.org/10.1063/5.0054894> .
17. *Bianca-Gabriela Antofie, Marius-Ciprian Larco, Teodor Lucian Grigorie* // Modern Applications of Wavelet Transform. / Ed by Srinivasan Ramakrishnan. IntechOpen, 2023. P. 85. <https://doi.org/10.5772/intechopen.1003199>.

ПОДПИСИ К РИСУНКАМ

- Рис. 1.** Схемы установки системы Interacoustics, Otometrics и Synapsys.
- Рис. 2.** Общая схема постановки эксперимента по записи данных vHIT-теста на смартфон: L – расстояние от оси штатива до головы пациента, H – расстояние от пола до оптической оси объектива камеры.
- Рис. 3.** Схема расчета угла рыскания α .
- Рис. 4.** Рефлектомограмма, записанная при положении головы P_2 .
- Рис. 5.** Результаты синхронизации данных с камеры смартфона и IMU-сенсоров: **а** – поворот головы влево, **б** – поворот головы вправо, **в** – нулевое положение. На кадре видна черная повязка, на которой крепится смартфон с IMU-сенсорами.
- Рис. 6.** Преобразование данных на выходе МО; подходит для DHPE, HPE, LHPE.
- Рис. 7.** Положение осей x , y , z в системе координат головы пациента.
- Рис. 8.** Гистограмма распределения коэффициентов искажения.
- Рис. 9.** Кривые опорного и искаженного сигналов гироскопа IMU по оси x .
- Рис. 10.** График зависимости взвешенных MSE от номера моды, полученной алгоритмом VMD, примененным к выходным сигналам по моделям МО: DHPE, LHPE, HPE по оси x .
- Рис. 11.** Сигналы DHPE, LHPE, HPE (после применения алгоритма VMD, выбраны оптимальные IMF) и гироскопа IMU по оси x .
- Рис. 12.** Графики зависимости взвешенных MSE от номера IMF, полученной алгоритмом EMD, примененным к выходным сигналам по моделям МО: DHPE, LHPE, HPE по оси x .
- Рис. 13.** Сигналы DHPE, LHPE, HPE (после применения алгоритма EMD, оптимальная IMF №2) и гироскопа IMU по оси x .
- Рис. 14.** Схема обработки сигнала с помощью вейвлет-преобразования.
- Рис. 15.** Сигналы LHPE и HPE (после применения алгоритма вейвлет-преобразования) с наименьшей взвешенной MSE и гироскопа IMU по оси x .
- Рис. 16.** Сигналы DHPE, LHPE, HPE (после фильтрации Баттервортса) и гироскопа IMU по оси x .
- Рис. 17.** Значения взвешенных ошибок MSE по оси x .
- Рис. 18.** Графики зависимости взвешенных MSE от изменения частоты среза фильтра Баттервортса по оси x .

Рис. 19. Сигналы DHPE, LHPE, HPE (после фильтрации Баттервортса с оптимальными частотами среза) и гироскопа IMU по оси x .

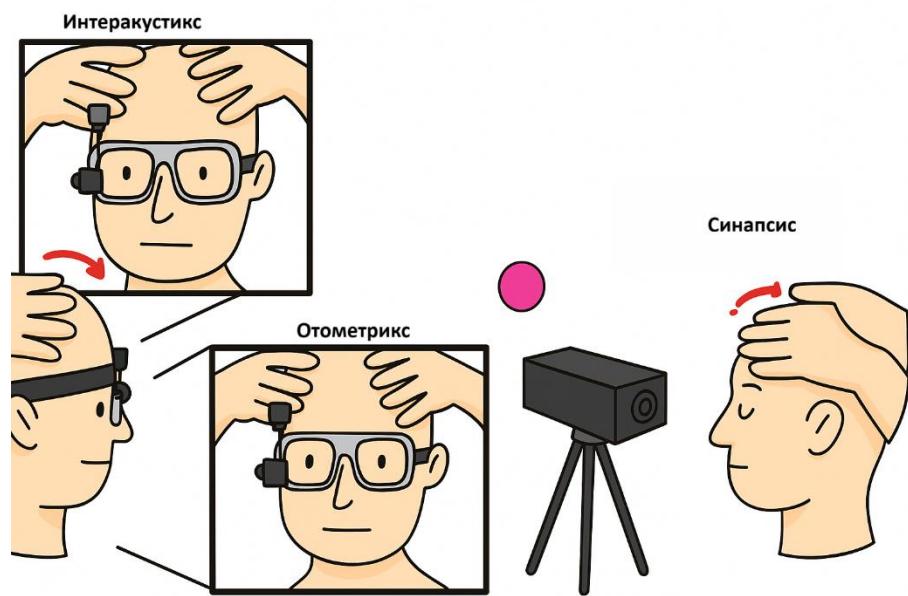


Рис.1.

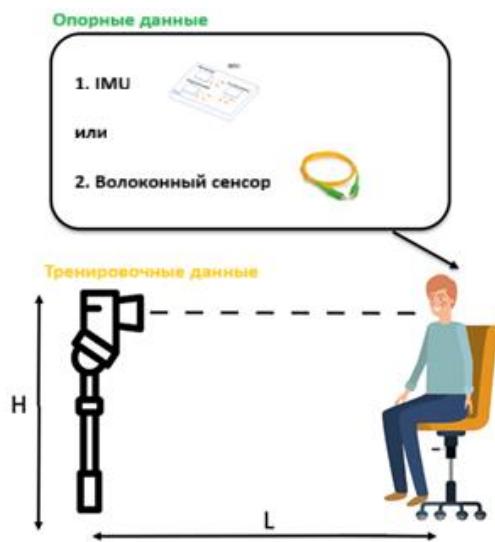


Рис.2.

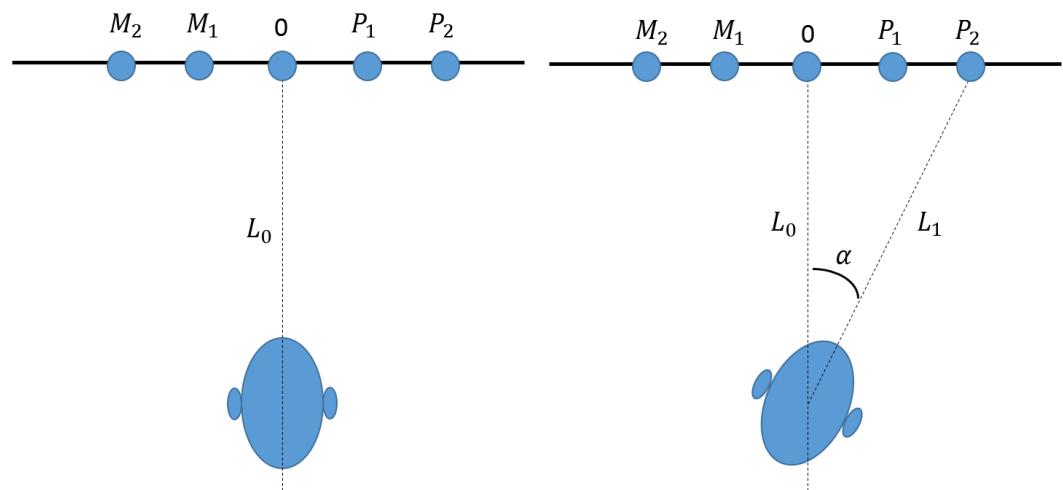


Рис.3.

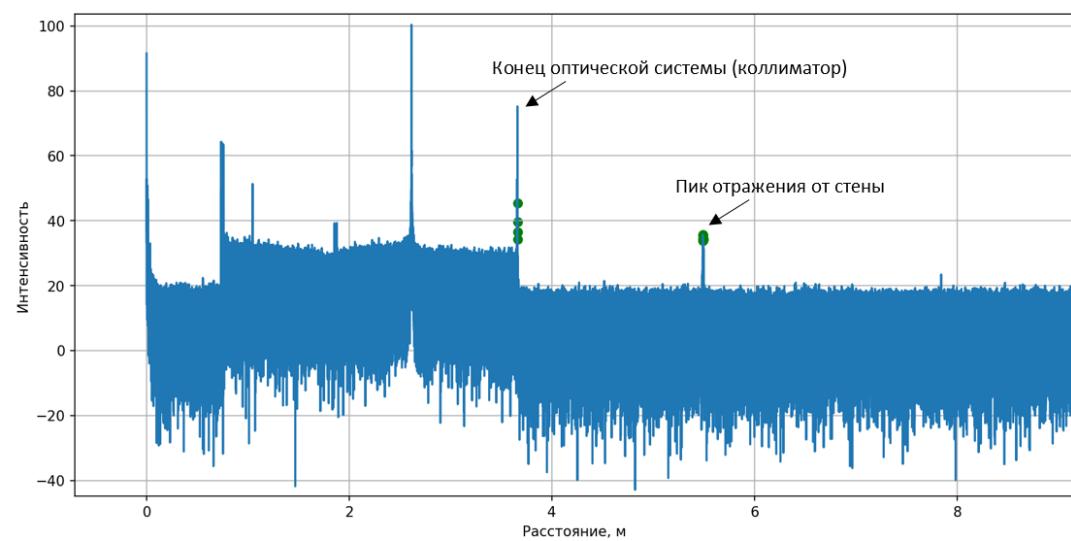


Рис.4.

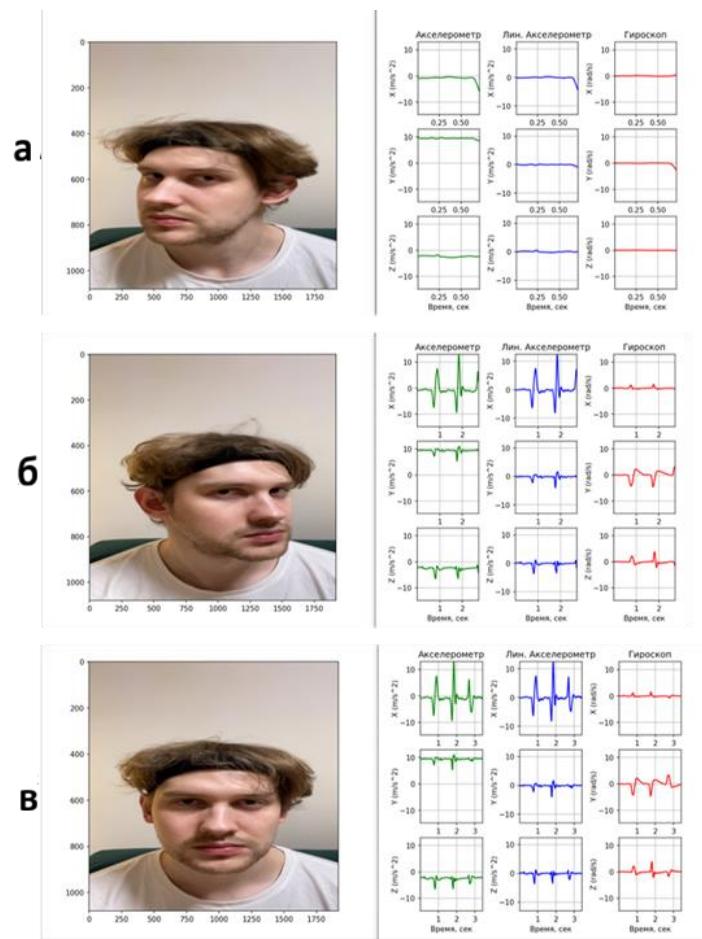


Рис.5.

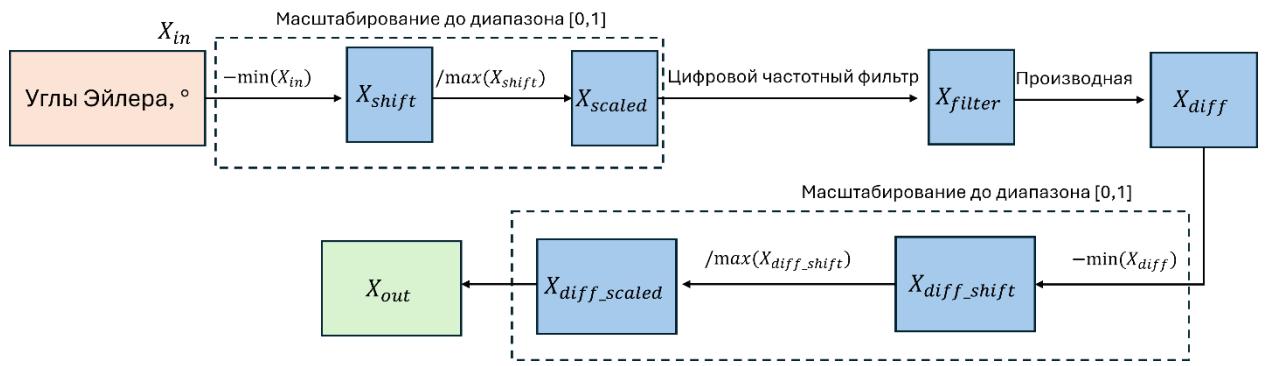


Рис.6.

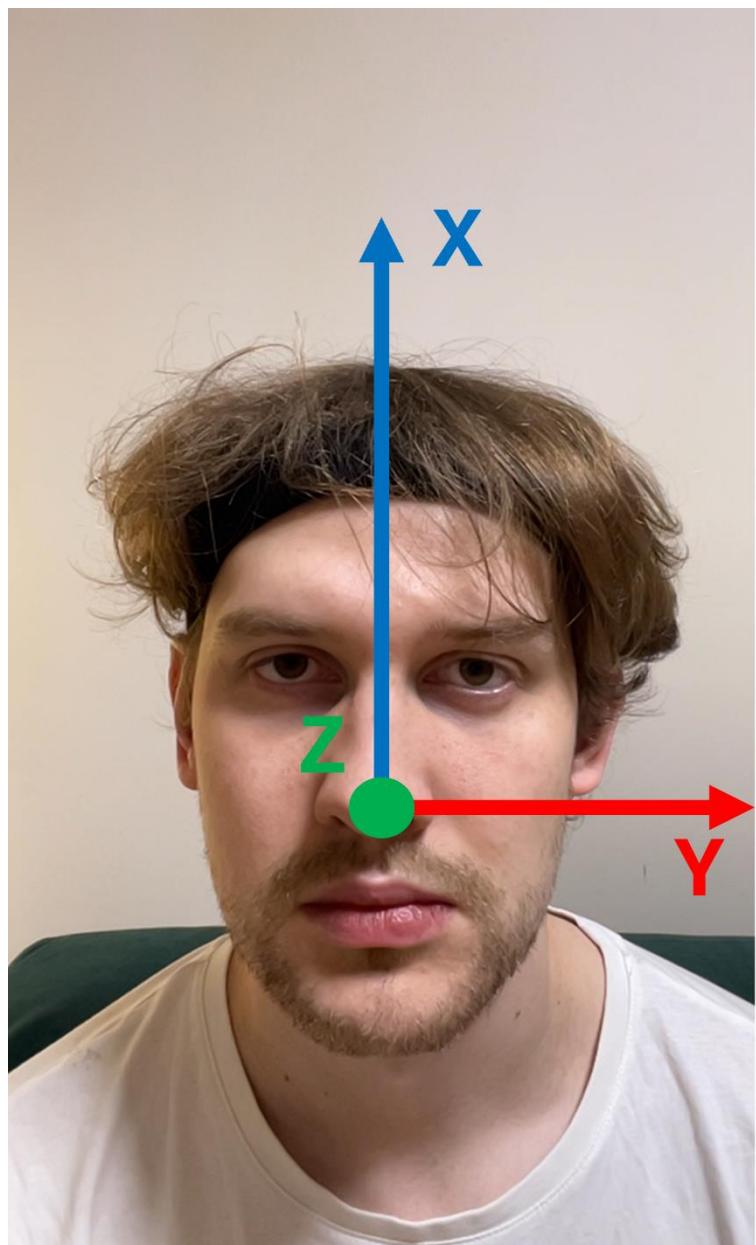


Рис.7.

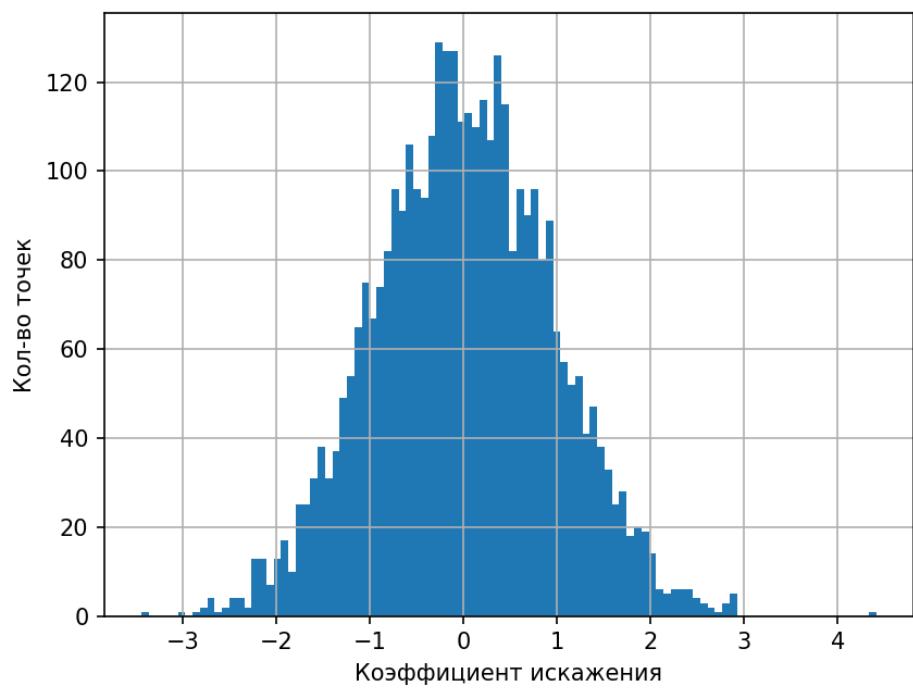


Рис.8.

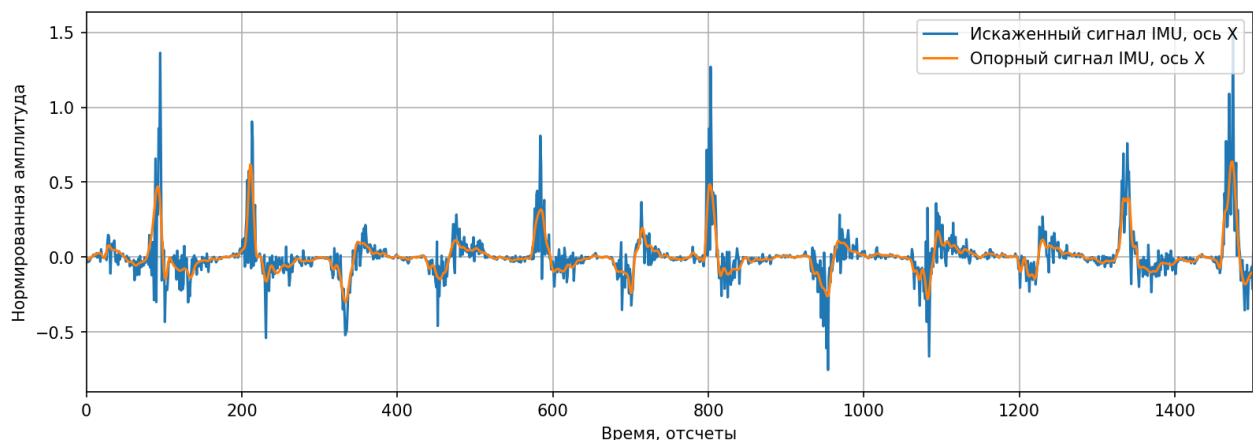


Рис.9.

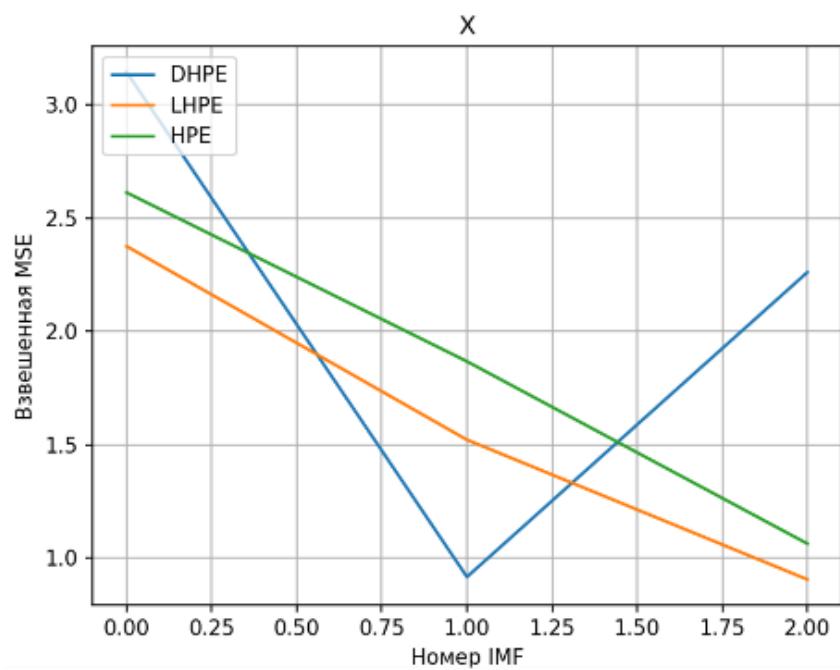


Рис.10.

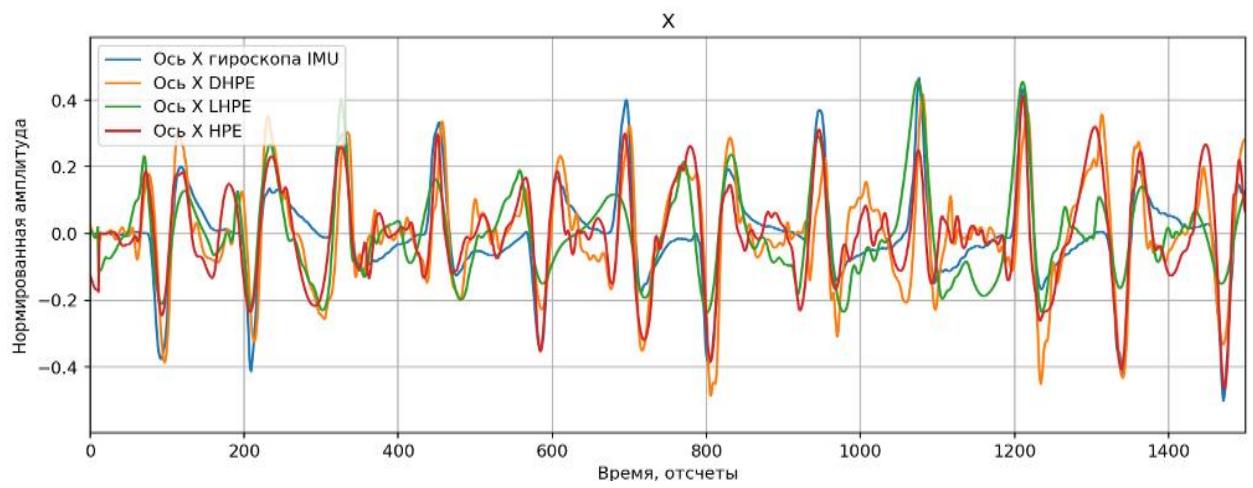


Рис.11.

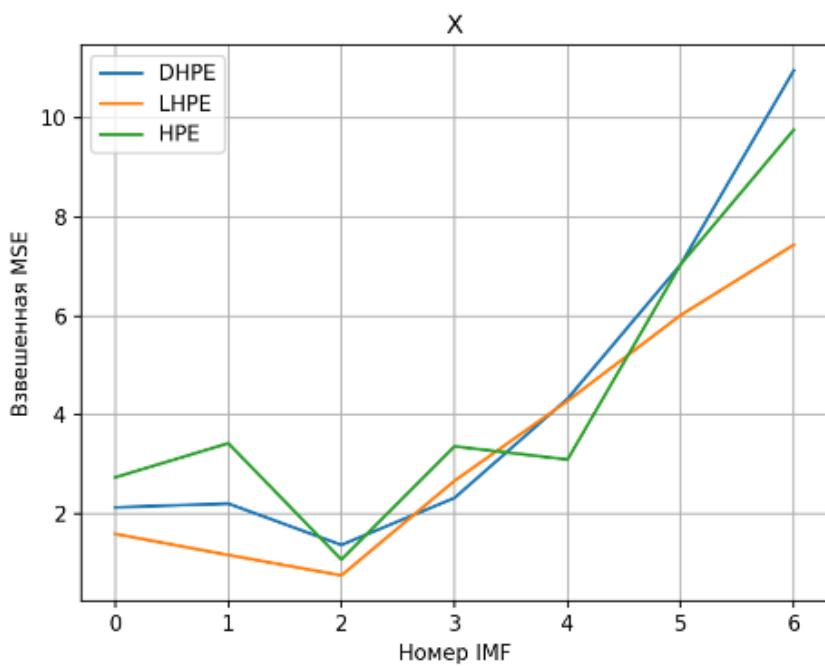


Рис.12.

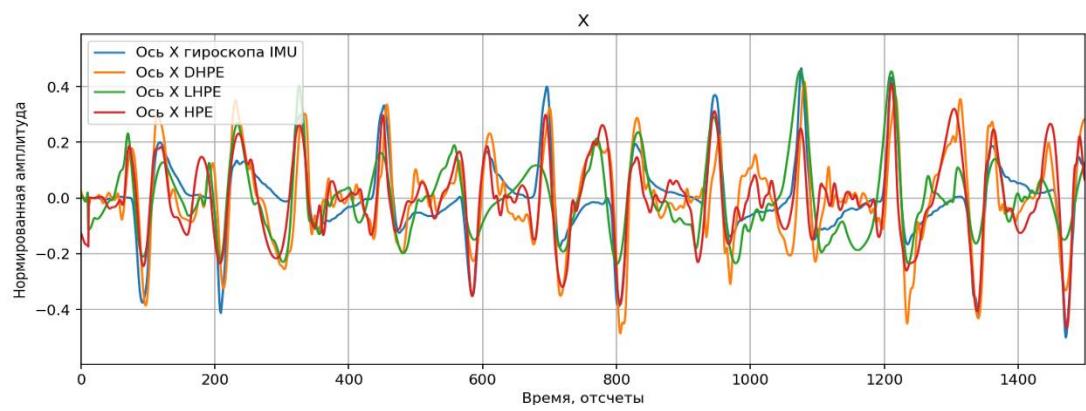


Рис.13.

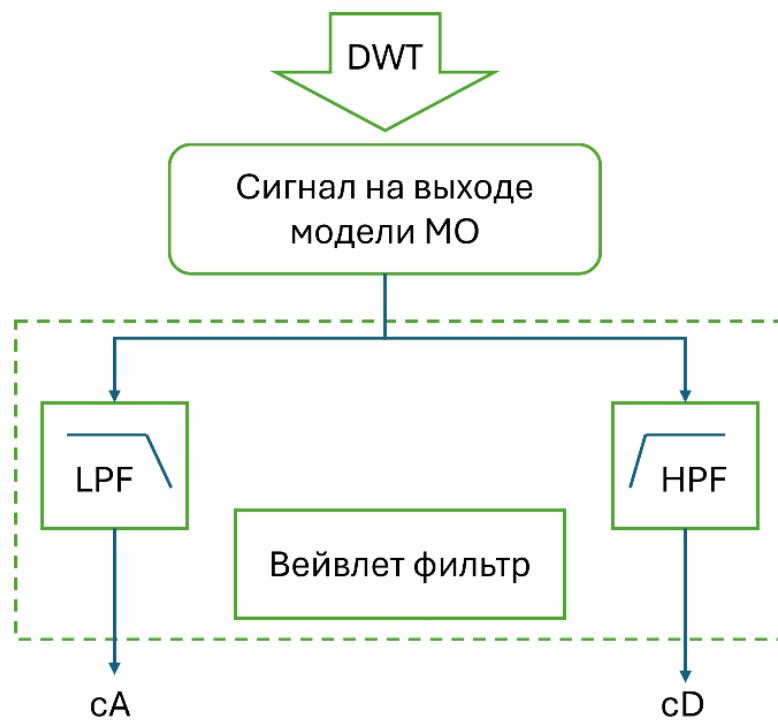


Рис.14.

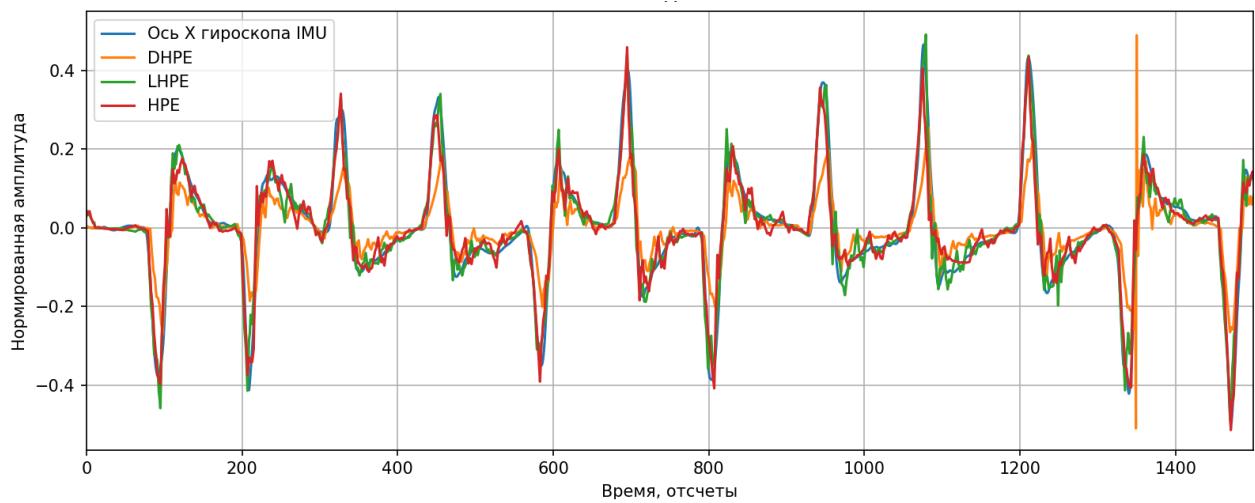


Рис.15.

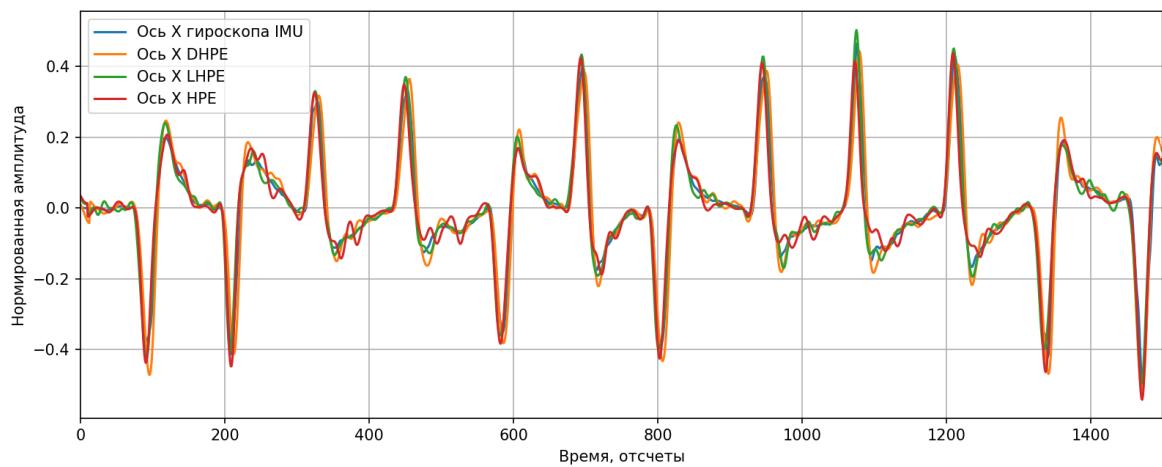


Рис.16.

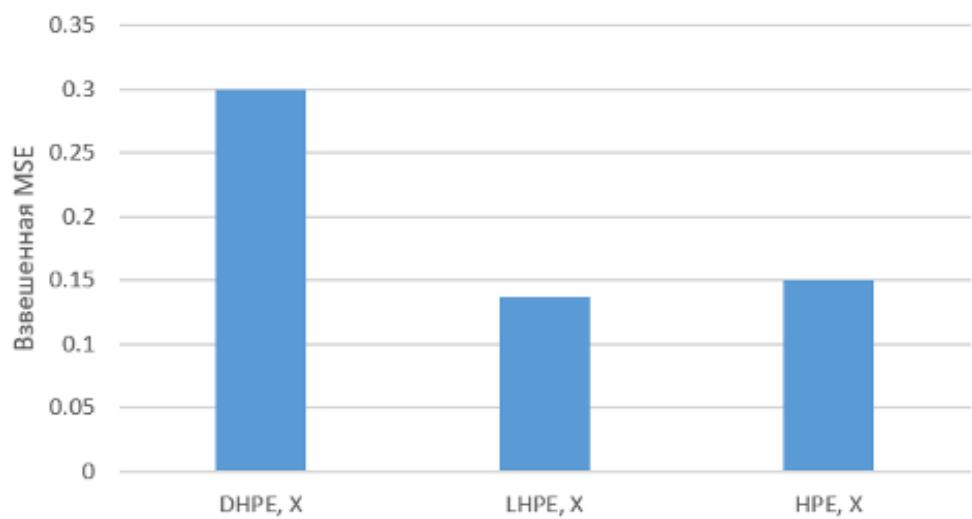


Рис.17.

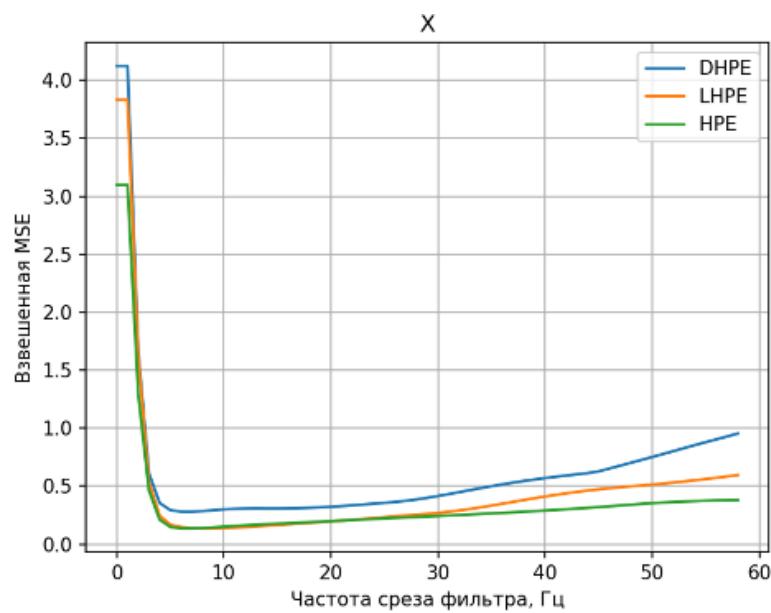


Рис.18.

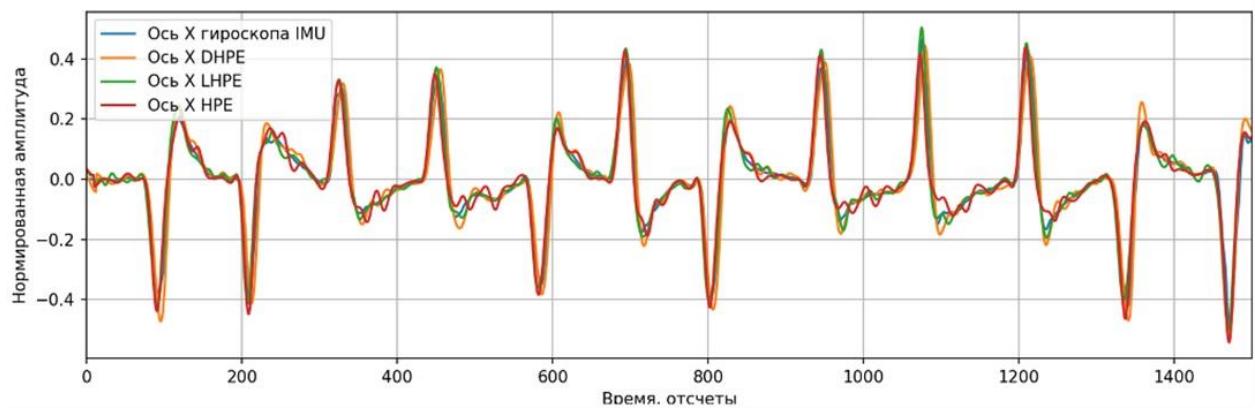


Рис.19.