



Московский государственный университет имени М. В. Ломоносова

Факультет вычислительной математики и кибернетики

Кафедра математических методов прогнозирования

Илларионова Светлана Владимировна

Нейрокомпьютерный интерфейс в задаче распознавания ошибок в трёхмерном видео

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель:

к.ф-м.н., доцент

С. И. Гуров

Москва, 2017

Содержание

1. Введение	3
2. Постановка задачи	5
2.1. Неформальная постановка задачи	6
2.2. Формализация постановки задачи	6
3. Обзор предметной области и существующих решений	8
3.1. Об интенсивности ошибок	11
3.2. Программное обеспечение для детектирования ошибок 3D	11
4. Построение решения задачи	11
4.1. Аппаратура для съёма ЭЭГ	13
4.2. Подготовка видео последовательности для проведения эксперимента	14
5. Описание практической части	17
5.1. Проведение эксперимента	17
5.2. Предобработка данных	19
5.3. Предобработка данных. Слуховая кора	20
5.4. Предобработка данных. Моргания	21
5.5. Выделение признаков из предобработанных данных	22
5.6. Частоты	24
5.7. Построение классификатора	25
6. Полученные результаты	25
7. Заключение	31
8. Список литературы	33

В данной работе рассматривается влияние ошибок трехмерного видео на зрителя. До этого проводились исследования о степени усталости человека от просмотра некачественного видео с помощью анкетирования, что является субъективной оценкой. Предложен метод, использующий данные электроэнцефалограммы. Проведены эксперименты для получения достаточного набора данных. Реализована программная часть, позволяющая осуществить предобработку данных, основываясь на биологической составляющей решаемой задачи. Рассмотрены различные ритмы головного мозга, на которых велось обучение классификаторов. Выявлены наиболее значимые для восприятия человеком ошибки.

1. Введение

Нейрокомпьютерные интерфейсы применяются для решения различных задач. Одна из таких задач связана с трёхмерным видео. При создании сконвертированного либо снятого фильма в формате 3-D могут возникать искажения. Объёмного эффекта в видео позволяет достигнуть то, что на каждый глаз подается своё изображение, отличающееся на определенный угол. При съёмке такого фильма на специальный штатив устанавливаются две камеры, которые аналогично глазам зрителя «видят» объект под своим углом. Но при съёмках, например, при движении штатива, этот угол может измениться. Это полученное отличие в ракурсах уже не будет направлено на создание эффекта объёма, а будет просто производственной ошибкой. Подобные ошибки, возможно, не будут бросаться в глаза зрителю в кинотеатре, то есть он не сможет сказать, что на конкретной секунде одна из камер сместилась. Но, в целом, после просмотра такого видео с некачественными фрагментами мозг человека будет вынужден подстраиваться под неестественные для глаз задачи обработки увиденного. Это может вызвать головную боль у зрителя.

Такие ошибки можно находить в видео. Но зачем это нужно? Не все несоответствия в изображениях, подаваемых на левый и правый глаз, обязательно должны привести к ухудшению самочувствия. Ошибки могут отличаться по своему типу и по интенсивности искажений. Некоторые ошибки способствуют резкой и сильной усталости от просмотра, а другие особо не влияют на зрителя. Но их исправление — трудозатратный и дорогой

Рис. 1. Пример 3D видео. Верхнее изображение предназначено для левого глаза, нижнее изображение — для правого глаза. Выделен пример отличия двух изображений для достижения эффекта объема.



процесс, для осуществления которого нужны веские причины. Такими вескими причинами может быть информация об усталости зрителя от просмотра видео и ранжирование различных ошибок. Имея такие данные производитель фильма может быть уверен в том, что дополнительные средства, потраченные на корректировку фильма, действительно потрачены с пользой: у зрителя не заболит голова, он пойдет вновь на фильм в формате 3-D.

Имеются реализации программного обеспечения поиска ошибок, как несоответствий

двух изображений, подаваемых в один момент времени. Выделены и изучаются определенные типы ошибок. Другой задачей является анализ этих ошибок по влиянию на самочувствие зрителя. Здесь недостаточно только набора самих видео, нужна также реакция зрителя.

Выделим следующие плюсы и минусы данной темы:

Плюсы

- Актуальность, обусловленная ростом процента фильмов в 3D.
- Есть возможность проведения субъективных сравнений с помощью анкетирования.

Минусы

- Нет готовых наборов данных для анализа.
- Сложность подготовки тестовых видеопоследовательностей, то есть ошибки добавляются самостоятельно в исходные видео.
- Нет готового алгоритма анализа сигналов мозга, распознающих усталость

2. Постановка задачи

- **Стереofilm** — видеоряд, использующий два отдельных изображения, позволяющих достичь стереоэффекта.
- **Стереопара** — пара плоских изображений одного и того же содержания в один момент времени, имеющая различия между изображениями, призванные создать эффект объёма.
- **Искажения, ошибки 3D видео** — несоответствие между видеорядами или кадрами стереофильма, которое не имеет отношения к естественному горизонтальному смещению видеокамер, обеспечивающему стереоэффект.
- **Усталость** — в данной работе негативное влияние ошибок 3D видео, то есть зрительных образов, в результате которых испытуемый ощущает дискомфорт, связанный с проявлениями головной боли (дополнительно явление усталости, рассматриваемое в работе, будет описано ниже с точки зрения работы мозга человека).
- **Электроэнцефалограмма (ЭЭГ)** — графическое изображение сложного колебательного электрического процесса, который регистрируется при помощи специального прибора, и является отображением активности мозга.

2.1. Неформальная постановка задачи

Необходимо научиться определять степень усталости зрителя при просмотре некачественного видео по средствам анализа электроэнцефалограммы (далее ЭЭГ).

Рассматриваются заданные ошибки 3D разной интенсивности, где множество ошибок содержит следующие типы:

- цветовой
- поворот
- временной сдвиг

Интенсивность искажений определяется по шкале с единичным шагом от 0 до 4, где 0 соответствует отсутствию ошибки, а 4 соответствует наиболее сильному искажению.

В данной работе не рассматривается влияние возможных других дефектов трёхмерного видео, и анализируются только вышеперечисленные артефакты. Для проведения исследования используется прибор Ерос Emotiv для записи сигналов мозга.

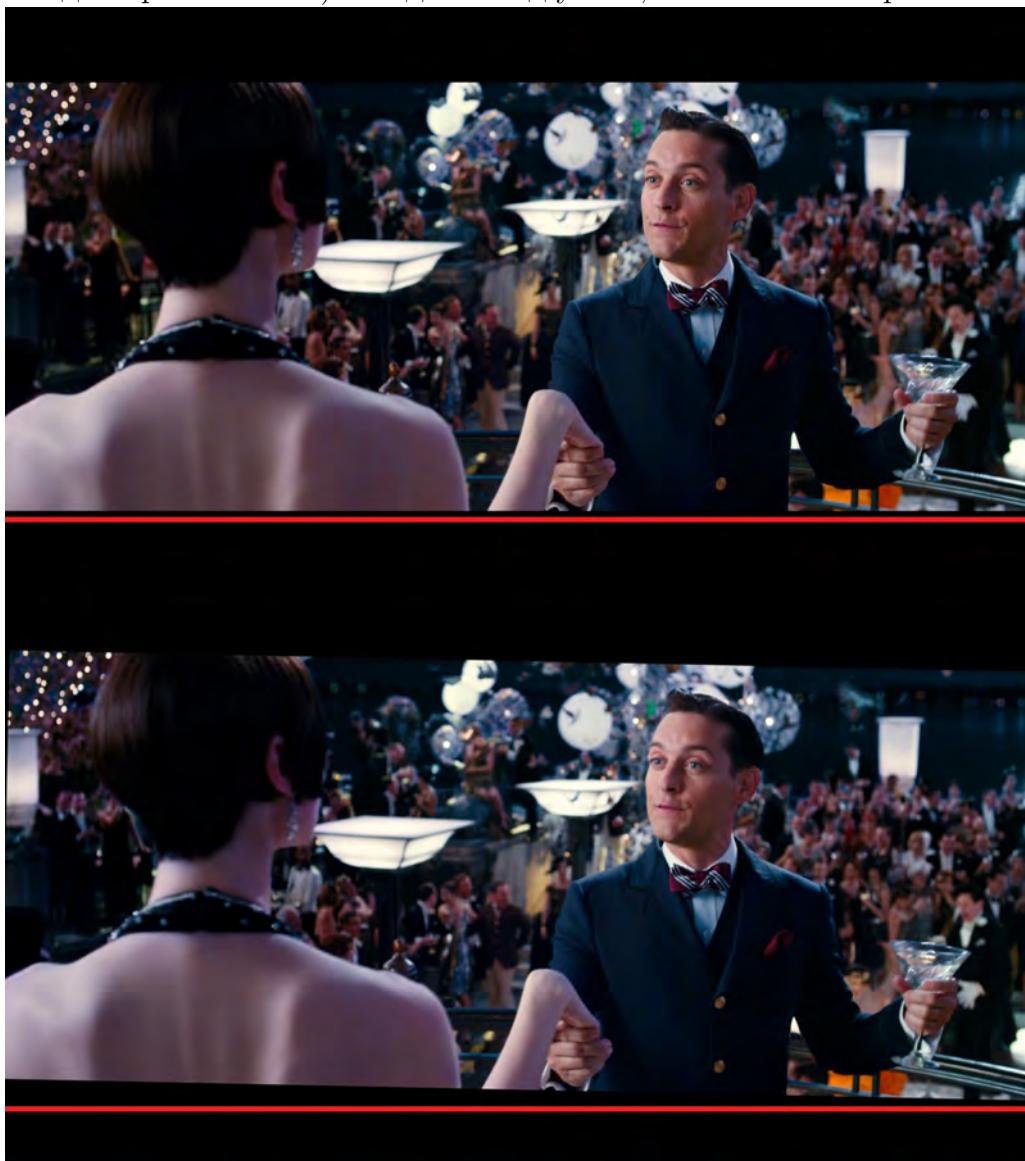
Данную проблему можно решать путем анкетирования, но возникает фактор необъективности оценки человека своего самочувствия. Так, нельзя сказать, что усталость на моменты заполнения анкеты будет вызвана именно просмотром последнего фрагмента видео. Возможна ситуация, когда человек после просмотра болезненного участка видео, не почувствует сразу усталость, а заметит существенное изменение в своем состоянии только в момент просмотра нейтрального (свободного от ошибок) фрагмента фильма. К тому же представляется сложным с полной уверенностью оценивать своё состояние даже по 10 балльной шкале. Однако, мозг при получении зрительного стимула обрабатывает его непосредственно в момент возникновения, что позволяет зафиксировать болезненный момент 3D видео без задержки.

2.2. Формализация постановки задачи

Необходимо решить задачи:

- 1) Разработка решения задачи выделения из ЭЭГ признаков, отвечающих за усталость при внешнем воздействии на человека при просмотре некачественного стерео изображения.

Рис. 2. Поворот. По красным параллельным линиям видно, что нижнее изображение (предназначенное для правого глаза) находится под углом, отличным от верхнего



- 2) Разработка и проведение эксперимента. Сбор необходимых данных для анализа. Проведение анкетирования для возможности реализации субъективного сравнения полученных результатов.
- 3) Проведение предобработки данных для повышения качества работы классификатора. Построение алгоритма, позволяющего по данным ЭЭГ предсказывать болезненность видео.

Рис. 3. Цвет



3. Обзор предметной области и существующих решений

Данная задача не решалась прежде рассматриваемым в работе методом, использующим нейрокомпьютерный интерфейс.

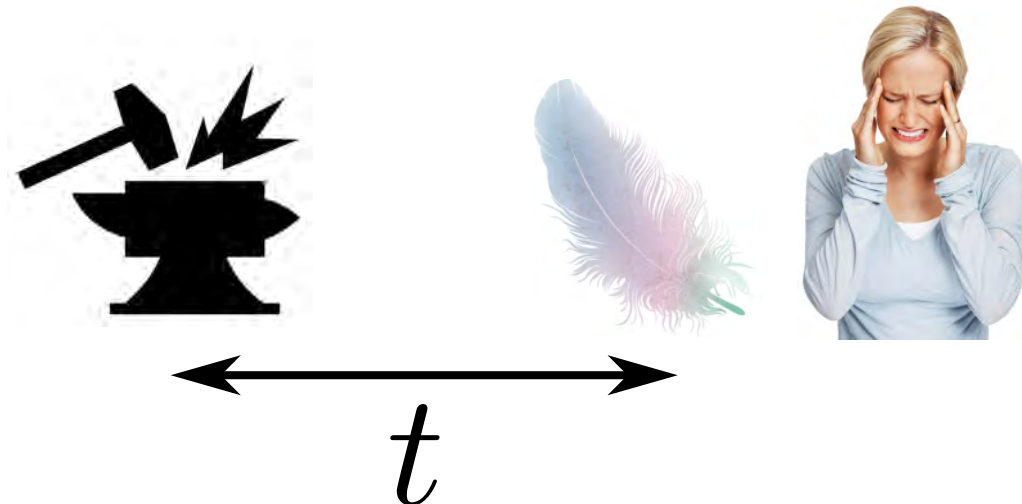
В статье «Emotion recognition using brain activity» рассматривается выявление эмоциональных состояний испытуемого с помощью прибора для измерения ЭЭГ [15]. Автором в первую очередь даётся определение термина «эмоции», как сложный комплекс взаимодействий между субъективными и объективными факторами, управляемый нервной/гормональной системами. Для анализа с помощью нейрокомпьютерного интерфейса автор разбивает эмоции на две составляющие, которые способен фиксировать прибор.

Рис. 4. Временной сдвиг



Первое это градация от негативного к позитивному, второе это степень эмоциональной насыщенности (от спокойного состояния к возбужденному). Большое внимание в работе уделяется проведению экспериментов для получения данных, на которых в дальнейшем происходит обучение классификатора. В качестве внешних эффектов, воздействующих на испытуемого, выбраны фрагменты видео фильмов. Каждый фрагмент фильма должен соответствовать определённому состоянию зрителя: радость, грусть, испуг и т.д. При выборе фильмов исследователи ориентировались на рейтинг кинематографа по разным жанрам, чтобы использовать наиболее репрезентативные для каждой категории. Далее производилась предобработка полученных сигналов, удаление посторонних артефактов.

Рис. 5. Человек может почувствовать дискомфорт от просмотра болезненного фрагмента спустя некоторое время t , при просмотре фрагмента без ошибок (либо с незначительными искажениями)



На уже отчищенных данных было обучено несколько классификаторов машинного обучения и произведено сравнение результатов их работы с целью выявления наилучшего метода решения поставленной задачи. Наиболее точные результаты были получены при использовании SVM классификатора.

В книге [12] освещаются проблемы, возникающие при работе с ЭЭГ, и возможные подходы к их решению. Имеется подробное описание предметной области, то есть работы мозга в разных условиях. Производится разбор и характеристика сигналов мозга по частотному диапазону.

Предложен подход к изучению влияния ошибок трехмерного видео с помощью механизма содружественного саккадического движения оптических осей глаз. Показано, что критическое значение экранного параллакса, при котором возникает эффект двоения стереоскопических изображений, зависит от их структуры, размера, а также расстояния наблюдения и может изменяться в широких пределах [16]. В проводимых экспериментах испытуемым предлагалось оценивать значение, при котором изображение начинало «двоиться» в глазах.

3.1. Об интенсивности ошибок

В данной работе рассматривалась интенсивность ошибок по равномерной сетке. В реальных фильмах интенсивность может меняться непрерывно. Однако:

- в экспериментальной видео последовательности нельзя рассмотреть случаи для непрерывного изменения ошибок (даже если максимально уменьшить шаг изменения интенсивности ошибок, то есть приблизить к непрерывному случаю, то время необходимое для записи эксперимента с одним испытуемым неприемлемо возрастет). Время проведения эксперимента ограничено временем высыхания датчиков (сигнал становится слабее) и тем фактом, что накопленная усталость с некоторого момента сделает полученные данные нерелевантными
- из ранне проведенных исследований известны рамки, в которых в среднем меняется интенсивность ошибок в фильмах.

Таким образом, достаточно рассмотреть определенный отрезок интенсивности каждой ошибки (например, интенсивность ошибки, связанной с цветовым искажением, не меняется в такой степени, что один глаз видит изображение белое, а другой черное). При выборе интервала изменения интенсивности учитывались результаты работы [19].

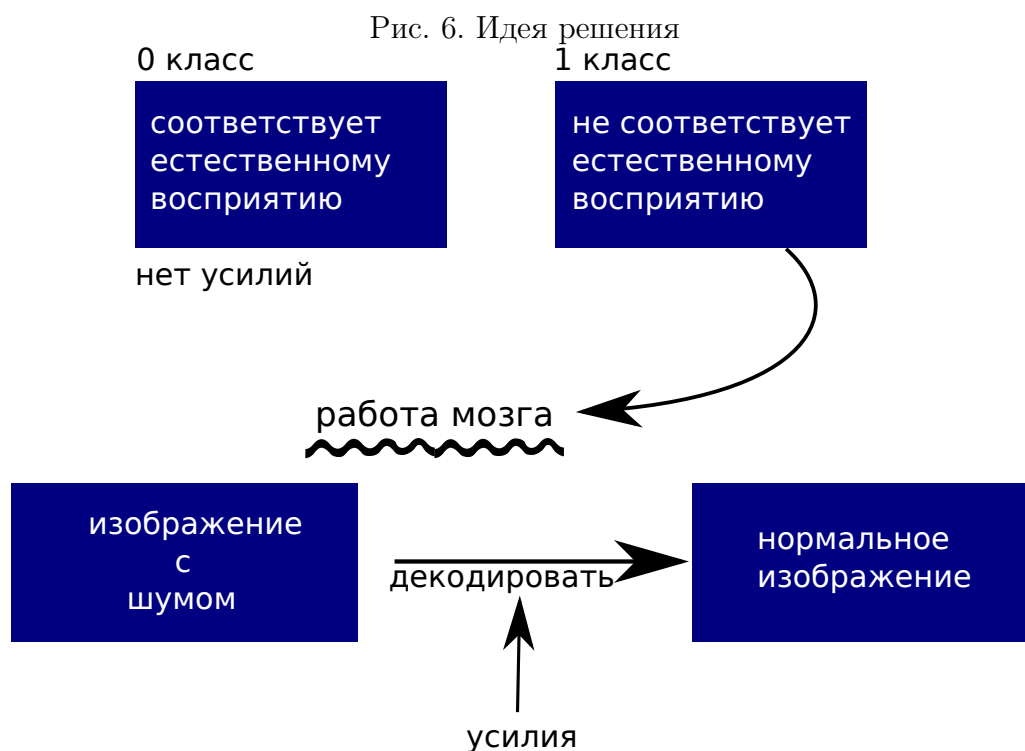
Получили для каждой ошибки свой диапазон изменения интенсивность. Разобьем этот диапазон на 4 равные части. Отметим здесь, что в работе не утверждается, что одинаковые интенсивности для разных ошибок равноценно болезненны.

3.2. Программное обеспечение для детектирования ошибок 3D

В данной работе для детектирования ошибок видео используется программное обеспечение, разработанное на кафедре АСВК ВМК МГУ лабораторией компьютерной графики и медиа (КГиМ). VQMT3D (Video Quality Measurement Tool) позволяет определять качество видео с использованием таких метрик как временной сдвиг, масштаб, цветовое искажение и т.д. [17]

4. Построение решения задачи

Предлагается следующий вариант решения поставленной задачи.



Зрительная система человека устроена так, что при анализе окружающего мира мозг человека прикладывает в среднем одинаковые усилия для зрительной обработки образов (речь не идет о следующих стадиях анализа этих образов). Однако, зрительная система не имеет опыта восприятия окружающего мира, содержащего искажения, подобные ошибкам ЗД, это вызывает головную боль и усталость [18]. Таким образом мы можем говорить о зрительных стимулах, соответствующих естественному восприятию, и зрительных стимулах, которые его нарушают. Известно, что если, поступающие на сетчатку глаза, изображения будут нарушать законы бинокулярного построения зрительной системы, мозг будет вынужден привести эти изображения («декодировать») к адекватному виду, который будет приемлемым для следующих стадий анализа изображения, происходящих уже вне зрительной коры. Эти усилия, затраченные мозгом, и превышающие работу на естественное восприятие, мы и будем в данной работе приравнивать усталости от просмотра (головной боли). Таким образом, чем больше усилий, тем больше отличие ЭЭГ от спокойного состояния.

Имеется набор заранее известных ошибок. Предполагаем, что данные ошибки вызывают ухудшение самочувствия зрителя при просмотре видео. Изменения в самочувствии фиксируются в энцефалограмме.

Проблемой является то, что изменения в ЭЭГ не могут быть очевидны невооружённому глазу при просмотре, например, графиков, полученных при записи. Это набор сигналов с определёнными амплитудами и частотами. Однако, данные ЭЭГ можно преобразовывать для дальнейшего их использования в качестве объектов, на которых учится классификатор.

**усилия, затраченные
на декодирование** **головная боль
(усталость от просмотра)**

Переходя к классификатору, нужно решить, что будем классифицировать, что является признаками и какие рассматриваются классы. Как было сказано выше, объектами будут данные, полученные преобразованием сигналов. Наша задача состоит в выделении признаков усталости, значит, разумно разделить объекты на два класса: усталость есть, усталости нет.

Тогда возникает вопрос: а как мы можем быть уверенными, что от каждой ошибки видео самочувствие зрителя ухудшается. Получается, что мы заведомо соглашаемся на то, что какая-то доля объектов будет нерелевантной, то есть не каждый фрагмент фильма приведёт к изучаемым изменениям. А значит, если объект не является представительным для своего класса, аналогичный ему объект тестовой выборки не будет отнесён классификатором с большой уверенностью к нужному классу.

Если, например, рассматривать объекты, представляющие ошибку видео i -го типа, и эта ошибка не влияет на усталость, то при классификации точность отнесения данных объектов к классу, соответствующему усталости, будет близка к случайной (классификатор не видит никаких закономерностей, на основании которых можно с точностью сказать, что человек устал).

Из сказанного выше получаем, что классификатору нужны два класса объектов. При этом классы должны отличаться только по качеству видео, где под качеством понимается наличие рассматриваемых ошибок 4-х типов с разной интенсивностью.

4.1. Аппаратура для съёма ЭЭГ

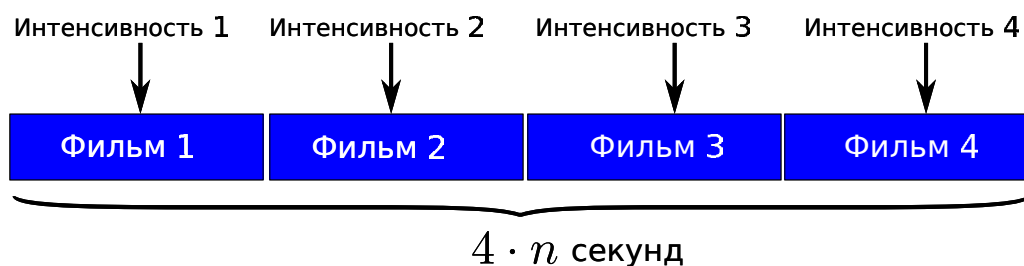
Для проведения экспериментов использовался прибор Eros Emotiv, оснащенный 14 датчиками. Он реализует неинвазивный (без непосредственного контакта с мозгом) метод

контроля работы мозга. Частота дескритизации данного прибора 128 Герц. Прибор находится в широкой продаже и используется, как для развлечения (под него разработано множество игр), так и для проведения научных исследований (в медицинских учреждениях и исследовательских лабораториях). Одним из его достоинств является его относительная доступность. Разработчики данной аппаратуры уверены, что в скором времени спрос на такой прибор возрастет. В продаже уже появились более мобильные и компактные приборы, которые можно использовать в повседневной жизни (например, для трекинга физического, эмоционального состояния, для использования людьми с ограниченными возможностями).

Перед каждым проведением эксперимента датчики прибора смачивались физраствором для лучшей проводимости. Далее прибор размещался определенным образом на голове испытуемого. Данные с прибора передавались в компьютер по беспроводному соединению.

4.2. Подготовка видео последовательности для проведения эксперимента

В качестве видео последовательностей выбраны фрагменты фильмов по 4 минуты из каждого фильма. Имеется 4 фильма. Используются фильмы снятые в 3D. В них добавляются ошибки в соответствии с приведённой снизу таблицей.



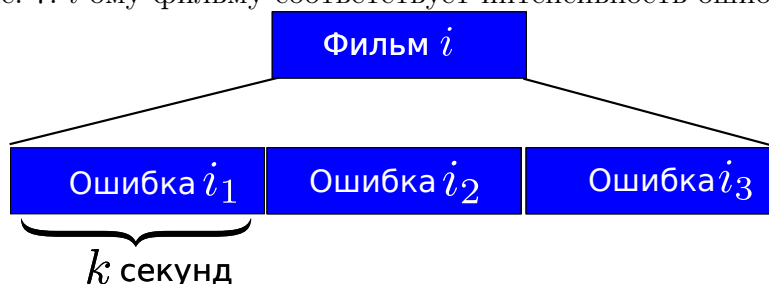
Каждый из этих артефактов (ошибок видео) вносится с определённой интенсивностью. Интенсивность определяется для каждого артефакта по своей равномерной дискретной сетке. При этом отметим, что в реальном 3D фильме интенсивность ошибки может варьироваться непрерывно. Число шагов в дискретной сетке интенсивности ошибки зададим для каждого вида ошибки одинаковым, равным 4.

В ходе эксперименты зрителю предлагается посмотреть отрывки из 4 фильмов продолжительностью 4 минуты каждый. Каждый фильм делится на 4 равные части. Каждая

часть отличается типом допущенной в ней ошибки 3D. При этом для каждого фильма выбирается своя интенсивность ошибки (на протяжении одного фильма интенсивность у всех ошибок одинаковая) и своя последовательность ошибок внутри фильма.

Описанное разбиение видеопоследовательности нужно для того, чтобы максимально избежать попадание в данные эксперимента ложной информации: например, в связи с тем что ошибки следуют в одном и том же порядке друг за другом, и возникает ошибочное предположение о сильном влиянии ошибки, на которой заканчивается последовательность ошибок одной интенсивности. То есть хотим, чтобы результаты не зависели от порядка следования ошибок.

Рис. 7. i -ому фильму соответствует интенсивность ошибки i



При выборе фильмов внимание уделялось тому, чтобы большая часть фильма была именно снята, а не сконвертирована в 3D.

Также важно, чтобы видео было максимально «однородными». Поясним это требование. Запись ЭЭГ позволяет фиксировать не только какие-либо воображаемые движения руки, например, но и изменение в настроении, эмоции. По этой причине нежелательным является использование в качестве некоторых фрагментов сцен, которые зритель может классифицировать, как «напряжённые» (например, драка), и сцен, которые зритель назвал бы «весёлыми/радостными» (например, красивые виды, радостные, смеющиеся лица). Также важна освещённость и яркость сцен в видео. Нежелательным будет наличие тёмных сцен (например, ночь) в одних видео и светлых сцен в других. Все эти требования выдвигаются как в сравнении всех фрагментов всех фильмов, так и внутри каждого фильма. Videopоследовательность была составлена в соответствии с вышесказанным.

Другим вопросом, возникающим при проведении эксперимента является необходимость добавления звуковой дорожки к видео. Эксперимент длится 20 минут, и на протяжении этого времени крайне желательным является то, что зритель не будет отводить глаза

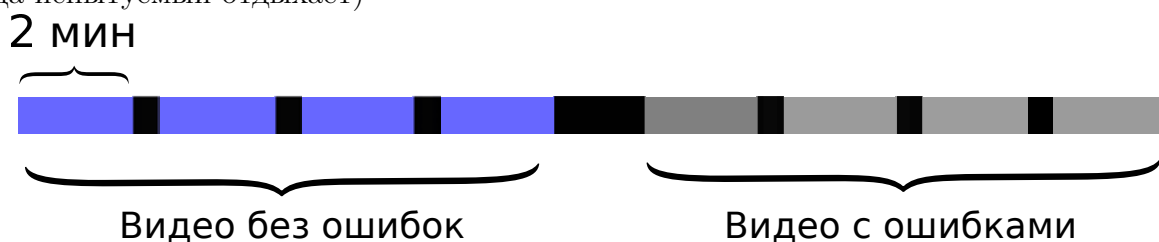
от экрана на продолжительное время, игнорируя тем самым изучаемый в эксперименте внешней эффект.

Эту проблему можно было бы решить, используя прибор для отслеживания движения глаз (Eye Tracker). Тогда можно узнать, действительно, ли испытуемый смотрел на экран, и подвергался влиянию 3D, и записанные данные ЭЭГ являются релевантными, либо были записаны данные, как испытуемый от скуки закрывает глаза и видео не смотрит.

Альтернативным решением такой проблемы является добавление звука к видео. 2 минуты одного сюжета (сцены выбраны из расчёта, что они должны быть интересны для просмотра, даже если человек не видела предыдущую часть фильма) со звуком смотрятся интереснее, чем без. Возможная сложность при таком проведении эксперимента — звук также записывается в ЭЭГ, и может исказить данные. Но решение этого вопроса будет подробнее описано в следующей части.

Итоговая последовательность видео выглядит так: 8 минут видео без искажений, 8 минут видео с искажениям. Между ними 1 минута, когда испытуемый может отдохнуть. Внутри каждой части между разными фильмами промежуток 10 секунд, когда экран чёрный, испытуемый также отдыхает.

Рис. 8. Разбиение видеоряда по качеству видео (чёрным обозначены части эксперимента, когда испытуемый отдыхает)



Такое разделение вызвано предположением о влиянии накопленной усталости: если перемешать качественное и некачественные фрагменты, то сложнее определить, чем вызвано изменение в самочувствие.

Интенсивность от 0 до 4, где 0 соответствует отсутствию ошибки

Искажения

A — цветовое (color), (цвет картинки, проецирующей на левый и правый глаз, различен)

В — поворот (rotate), (угол, под которым проецируется картинка на левый и правый глаз, различен)

С — временной сдвиг (temporal shift), (проецирование картинки на один глаз происходит с задержкой, относительно другого глаза)

	интенсивность	ошибка
<i>movie1</i> ₁₁	1	С
<i>movie1</i> ₁₂	1	А
<i>movie1</i> ₁₃	1	В
<i>movie2</i> ₁₀	2	А
<i>movie2</i> ₁₁	2	В
<i>movie2</i> ₁₂	2	С
<i>movie3</i> ₁₁	3	А
<i>movie3</i> ₁₂	3	В
<i>movie3</i> ₁₃	3	С
<i>movie4</i> ₁₀	4	В
<i>movie4</i> ₁₂	4	С
<i>movie4</i> ₁₃	4	А

5. Описание практической части

5.1. Проведение эксперимента

В эксперименты приняли участие 17 человек возраста от 20 до 22 лет. Каждому было предложено посмотреть в активных 3D очках 20 минут видео. В конце просмотра было проведено анкетирование с следующими вопросами:

- 1) Как часто вы посещаете кинотеатр, сеансы 3D?
- 2) Испытываете ли вы усталость/головную боль после просмотра фильма?
- 3) Как вы оцениваете своё самочувствие после просмотра первой части видео: 8 минут без искажений? По шкале от 1 до 10, где 1 — очень хорошо, 10 — очень плохо.

- 4) Как вы оцениваете своё самочувствие после просмотра второй части видео: 8 минуту с искажениями? По шкале от 1 до 10, где 1 — очень хорошо, 10 — очень плохо.
- 5) На что именно вы жалуетесь после просмотра видео?
- 6) На сколько интересно было для вас видео по шкале от 1 до 10, где 1 — очень скучно, 10 — очень интересно?

Также были добавлены некоторые замечания о ходе эксперимента.

При длительном действии адекватного стимула порог срабатывания данного рецептора повышается. Это явление называется адаптацией [11].

Это замечание в данной задаче можно интерпретировать следующим образом: люди, которые чаще посещают кинотеатр и 3D сеансы, более адаптированы к возникающим в видео ошибкам. Так, на пример, люди, занимающиеся анализом трёхмерного видео, и непосредственно ошибками видео, не испытывают головной боли и дискомфорта. Этот факт приводит к тому, что данные, выделенные из сигнала человека, часто смотрящего 3D, должны учитываться с меньшим весом, чем данные человека, глаза которого не адаптированы к восприятию искажений. Для этого и был добавлен в анкету первый пункт.

При рассматривании различных объектов внешнего мира глаза совершают быстрые (саккады) и медленные следящие движения. Благодаря медленным следящим движениям изображение движущихся объектов удерживается на fovea. При рассматривании хорошо структурированного изображения глаза совершают саккады, перемежающиеся с фиксацией взора. Если человек рассматривает изображение в течение некоторого времени, то запись перемещений глаза воспроизводит достаточно грубо контур и наиболее информативные детали рассматриваемого объекта. Например, при рассматривании лица особенно часто фиксируются рот и глаза [11].

Особый интерес представляет то, на сколько внимательным испытуемый был к видео. Движение глаз, как указано выше, можно отслеживать, что позволяет более детально анализировать влияние искажение в видео на зрителя. Это, например, можно реализовать с помощью айтрекера. Но при проведение данного эксперимента его использование было невозможно из-за необходимости использование 3D очков. Шестой вопрос был направлен на выявление того, на сколько внимателен испытуемый был к видео, так как если видео для него было неинтересным, то можно предположить, что человек не всё время смотрел на экран, а значит данные будут менее релевантными для анализа усталости от видео.

Перед проведением эксперимента испытуемого инструктировали о ходе эксперимента: нежелательны любые движения, так как они особо чётко записываются прибором. Также просили выключить телефон, чтобы уменьшить постороннее влияние на запись сигнала.

Вопросы	Средний результат по всем испытуемым
Усталость после первой части	2.31
Усталость после второй части	4.62
Интерес к видео	8.18

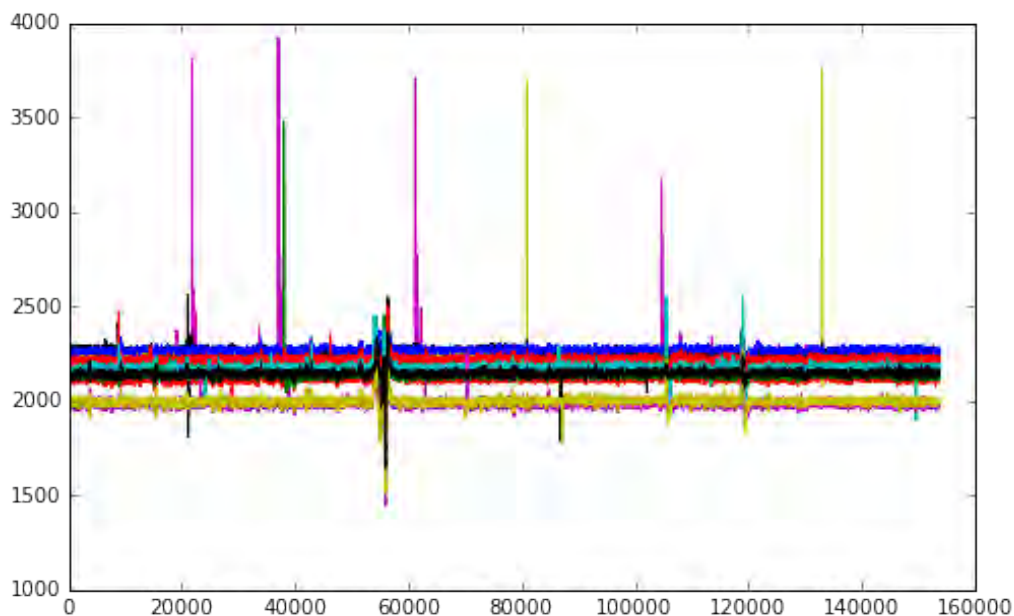
Основными характеристиками самочувствия после просмотра видео была: головная боль и усталость глаз.

5.2. Предобработка данных

Необработанные данные имеют следующий вид:

Рис. 9. Данные с 14 каналов без преобразований. Различными цветами выделены датчики.

По оси ординат — уровень сигнала, по оси абсцисс — сэмплы



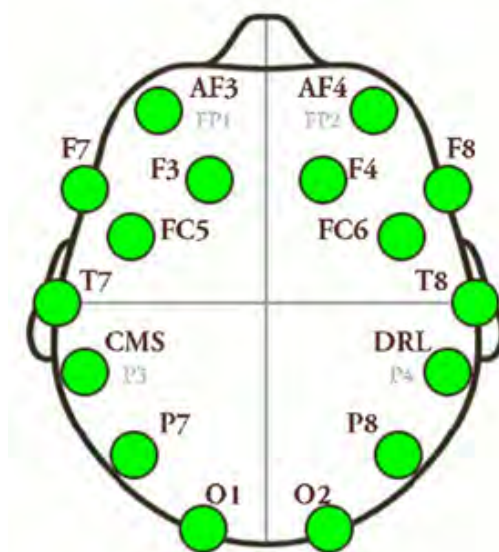
По горизонтальной оси отмечены сэмплы, по вертикальной оси — уровни сигналов. Разными цветами отмечены каналы датчиков.

На приборе для съёма сигналов расположено 14 датчиков, пронумерованных по международной системе размещения электродов «10-20»/: AF3, AF4, F7, F8, F3, F4, FC5, FC6,

T7, T8, P7, P8, O1, O2 (размещение датчиков приведено на рисунке 3). Каждый датчик выдаёт уровень сигнала. Признаки выделяются с помощью преобразования сигналов.

Дискретный сигнал — это последовательность уровней, то есть частоты. Дискретизация непрерывного сигнала производится за счёт измерения его каждый интервал времени. Дискретизируем с частотой 128 Герц. Выбор именно этой частоты обусловлен предустановками в драйвере для Ерос.

Рис. 10. Размещение датчиков



Референты — две точки, относительно которых производятся замеры при биполярного съёма. Референтами являются датчики *T7, T8*.

Шум — погрешности в записи ЭЭГ, возникающая из-за неинвазивности метода (зазор между датчиками и кожей), расстоянием между датчиками.

Получаемый при съеме ЭЭГ сигнал содержит в себе три вида данных: естественная деятельность мозга, шум, который накладывается на сигнал, реакция мозга на внешние стимулы, рассматриваемые в рамках решения задачи. Следовательно, для восстановления информации о влиянии на мозг внешних стимулов необходимо убрать шум.

Использование сырых данных не годиться для корректной работы классификатора.

5.3. Предобработка данных. Слуховая кора

На рисунке зелёным цветом выделена первичная слуховая кора. В системе размещения электродов датчики соответствующие этой зоне мозга — *CMS, DRL*. Данные с этих

датчиков для дальнейшего анализа не используются, чтобы уменьшить зашумленность. Действительно, при сравнение работы с ними и без них виден значительное улучшение при их исключении из рассмотрения.

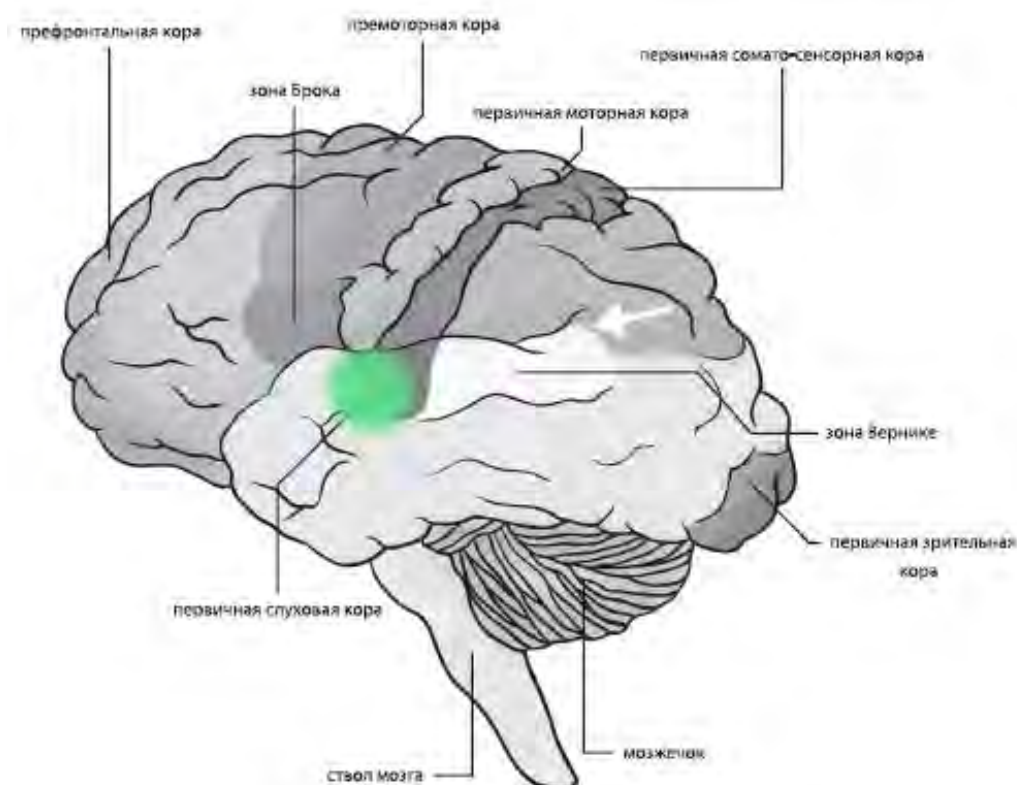
Рис. 11. Визуализация по 14 каналам данных, содержащих шум. По оси ординат — каналы, по оси абсцисс — сэмплы. Выделены шумовые фрагменты.



5.4. Предобработка данных. Моргания

Также из данных требуется исключить шум, связанный с морганием. Для этого используются данные с датчика *AE3*. К этому каналу применяется SVD (singular value decomposition) разложение. На выходе получаются свободные от артефактов моргания сэмплы ЭЭГ. Можно проследить, сколько таким образом сэмплов было удалено из данных. Данная функция реализована итеративным нахождением сэмплов, отклонение которых превышает 2.5 стандартных отклонения и последующим применением SVD. После SVD разложения вырезаются сэмплы с пиками, обнаруженными в ходе итеративного процесса. В результате применения SVD получаем матрицу меньшего ранга, которая ближе всего к исходной.

Рис. 12. Первичная слуховая кора



5.5. Выделение признаков из преобработанных данных

Обработка зрительных образов начинается воздействием на сетчатку глаза. Там зрительные сигналы преобразовываются в нервные импульсы и передаются для дальнейшей обработки в соответствующие части мозга. Для максимально эффективного использования данных, записанных с помощью 14-канального прибора, необходимо определить, какие датчики расположены непосредственно над зонами обработки зрительных сигналов.

Известно, что аксоны нейронов наружного колленчатого тела проецируются в кору затылочного полюса больших полушарий (поля 17, 18, 19). Высший этап обработки зрительных сигналов происходит в ассоциативных полях коры больших полушарий [13].

Обработка зрительных стимулов в ассоциативных полях коры. Исследование мозга низших обезьян показало, что анализ зрительной информации не завершается в полях зрительной коры (17, 18, 19). Было установлено, что от поля 17 начинаются пути (каналы), в которых производится дальнейшая обработка зрительных сигналов. Один путь, идущий в дорсальном направлении к полям теменной ассоциативной коры, участвует в формиро-

вании пространственного зрения, другой путь — вентральный (нижневисочная кора) — участвует в формировании предметного зрения [13].

Из этого следует, что для определения влияния зрительных стимулов на испытуемого недостаточно рассматривать только сигнал с датчиков, отвечающих зрительной коре. Активность разных участков мозга взаимосвязана, и требует комплексного анализа, по возможности, всех зон.

Другим вопросом является то, что неизвестно, какой участок мозга отвечает конкретно за усталость, и есть ли это какая-то отдельная ограниченная область, либо усталость от зрительных образов распространяется на всю кору.

Теорема Котельникова. *Если аналоговый сигнал не содержит в своём спектре частот выше F_{max} , то его можно идеально точно восстановить по дискретным отсчётам, взятым равномерно с частотой строго большей $2 \cdot F_{max}$.*

Из этой теоремы следует вывод, что для преобразования рассматривать частоты выше 63 Герц не имеет смысла, так как они будут определяться неточно.

После очистки данных от шумовых сигналов применяется Common spatial pattern (CSP). Главная его идея заключается в том, что производится сравнение объектов двух классов и подбираются веса, которые позволяют максимизировать расстояние между объектами двух противоположных классов. Метод подробно описан в работе [21].

$$W = \arg \max_w \frac{\|W X_1\|^2}{\|W X_2\|^2}$$

X_1, X_2 — матрицы, соответствующие двум классам. Размерность n на d , где n — число объектов, d — число признаков.

$$R_1 = \frac{X_1 X_1^T}{\text{tr}(X_1 X_1^T)} \quad R_2 = \frac{X_2 X_2^T}{\text{tr}(X_2 X_2^T)}$$

$$R = R_1 + R_2 = U_0 \Sigma U_0^T$$

Σ — диагональная матрица собственных значений. U_0 — матрица собственных векторов. R_1, R_2 — усредненная нормализованная ковариационная матрица.

$$P = \Sigma^{-1/2} U_0^T$$

$$S_1 = P R_1 P^T \quad S_2 = P R_2 P^T$$

$$S_1 = U\Sigma_1U^T \quad S_2 = U\Sigma_2U^T \quad \Sigma_1 + \Sigma_2 = I$$

$$W = U^T P$$

Далее данная матрица W используется для преобразования исходного сигнала.

5.6. Частоты

В ядрах таламуса действует генератор, посылающий ритмические импульсы по таламо-кортикальным волокнам к нейронами коры. Ритм может быть обусловлен либо пейсмейкерными свойствами самих клеток таламуса, либо активностью нейронных сетей с возбуждающими и тормозными синаптическими связями. Импульсы от таламуса вызывают ритмическое возникновение деполяризирующих синаптических потенциалов в апикальных дендритах пирамидных клеток коры. При этом в дендритах протекают токи, приводящие к генерации потенциалов действия. Одновременно возникают и петли внеклеточных токов, часть которых проходит обратно в непосредственной близости от дендритов. Чем более синхронно возбуждаются дендриты и чем плотнее они расположены, тем большая часть обратного внеклеточного тока будет протекать через окружающие ткани, включая черепную коробку и кожу головы. Зубец ЭЭГ соответствует разнице между величинами токов, протекающих в области электрода 1 и в области электрода 2. Если пирамидные клетки P_1 и P_2 возбуждаются одновременно, то возникающие при этом токи будут суммироваться и зубцы ЭЭГ будут выше.... Если же ритмы, навязываемые ядрами таламуса клеткам P_1 и P_2 , будут разными, то активность этих клеток будет несинхронной и их токи суммироваться не будут [13].

Например, альфа ритмы, которые свойственны состоянию спокойствия, ярко выражены при закрытых глазах, являются следствием синхронного возбуждения клеток. На энцефалограмме это отображается как высокие зубцы.

Разным уровням сознания соответствуют разные формы активности головного мозга [13].

При рассмотрении сигналов мозга человека выделяют активность следующих частот: альфа, бета, гама, тета ритмов. Данные волны характерны для различных состояний функционирования организма человека. Следовательно, важно понять, какие частоты являются репрезентативными для рассмотрения явления усталости.

- Тета волны 4 — 8 Герц. Состояние полудрема, неглубокий сон.

- Альфа волны 8 — 12 Герц. Возникают при закрытых глазах.
- Бета волны 14 — 40 Герц. Состояние бодрствования.

5.7. Построение классификатора

Класс объекта	Обучение	Тестирование
1	i -я ошибка заданной интенсивности	видео без i -й ошибки
0	часть видео без ошибок, соответствующая i ошибки	оставшееся видео

В качестве классификатора был использован Random Forest и xgboost.

Для каждого типа ошибок классификатор обучается отдельно. При этом для ошибки i -го типа в обучающую выборку входят все объекты, кроме объектов с рассматриваемым типом артефакта. В тестовой выборке содержатся объекты видео без ошибок, как представители одного класса, и объекты, соответствующие ошибкам i -го типа разной интенсивности. Таким образом классификатор обучается выделять признаки, указывающие на изменение состояние испытуемого из-за некачественного видео. Получаем для каждого типа ошибок свои результаты работы классификатора, по которым можно делать выводы о значимости ошибок, а так же для отдельных фрагментов видео определять уровень его влияние на зрителя.

Число признаков варьируется в зависимости от частот, на которых происходит обучение и составляет около 100 признаков. Для β -ритмов число признаков — 80, для базового частотного диапазона число признаков — 120. Число объектов после предобработки данных ЭЭГ одного испытуемого варьируется около 20000. После предобработки данных каждого испытуемого было удалено около 230 артефакта, соответствующих движению глаз.

6. Полученные результаты

Алгоритм запустим на данных, не использованных при обучении. Каждому столбцу графика соответствуют объекты выборки из одного фрагмента фильма продолжительностью 30 с. Соответственно, в каждом столбце равное число объектов. Высота столбца определяется как средняя уверенность классификатора в принадлежности объектов классу видео с ошибками. Черным цветом обозначается видео, в котором ошибок не было. Интенсивности ошибок расположены в соответствии с экспериментальной разбивкой видео.

Рис. 13. Обучение на временном сдвиге, полный частотный диапазон. Черным столбцам соответствуют фрагменты видео, где ошибок нет. Начиная с цветных столбцов (наличие ошибок), интенсивность каждых двух последующих столбцов возрастает на единицу.

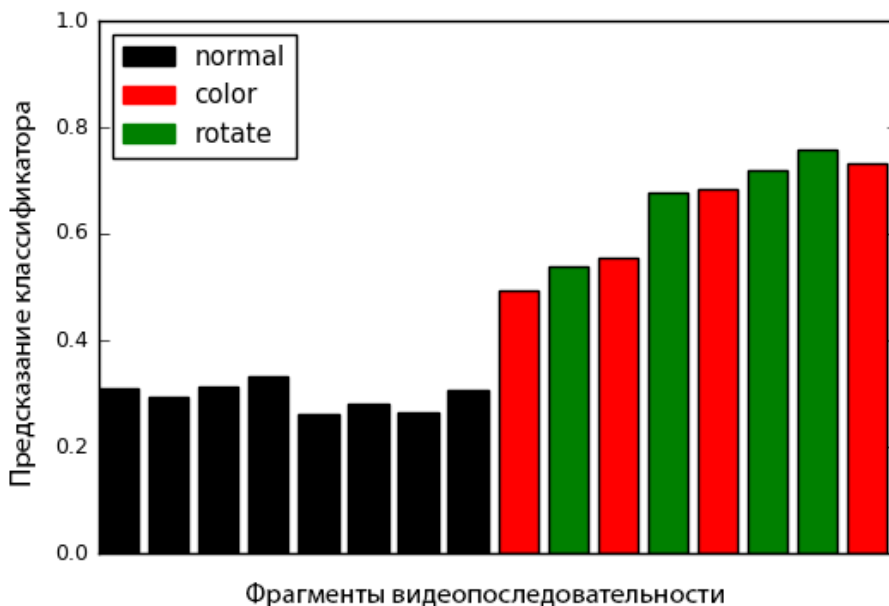
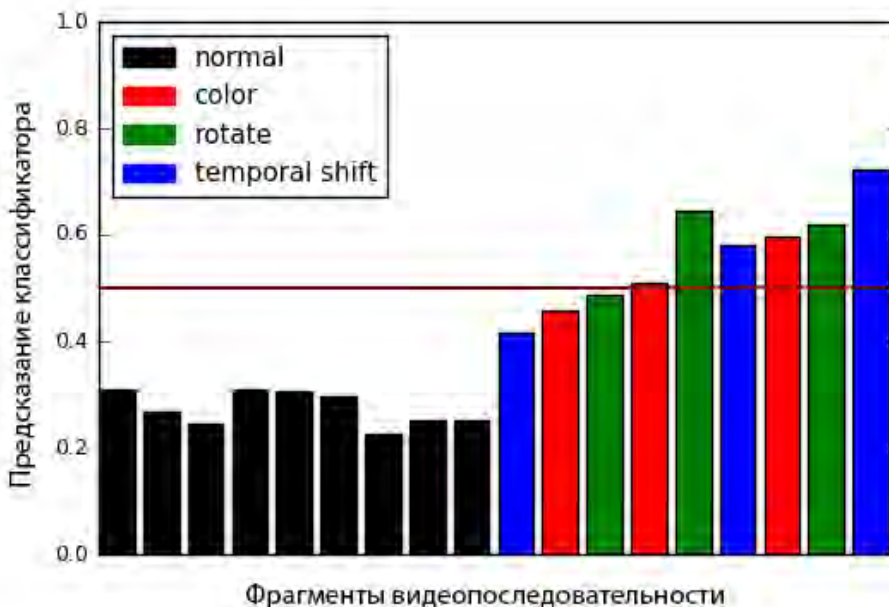


Рис. 14. Обучение на всех ошибках, полный частотный диапазон



$$precision = \frac{TP}{TP + FP}$$

Рис. 15. Обучение на всех ошибках, β -ритмы

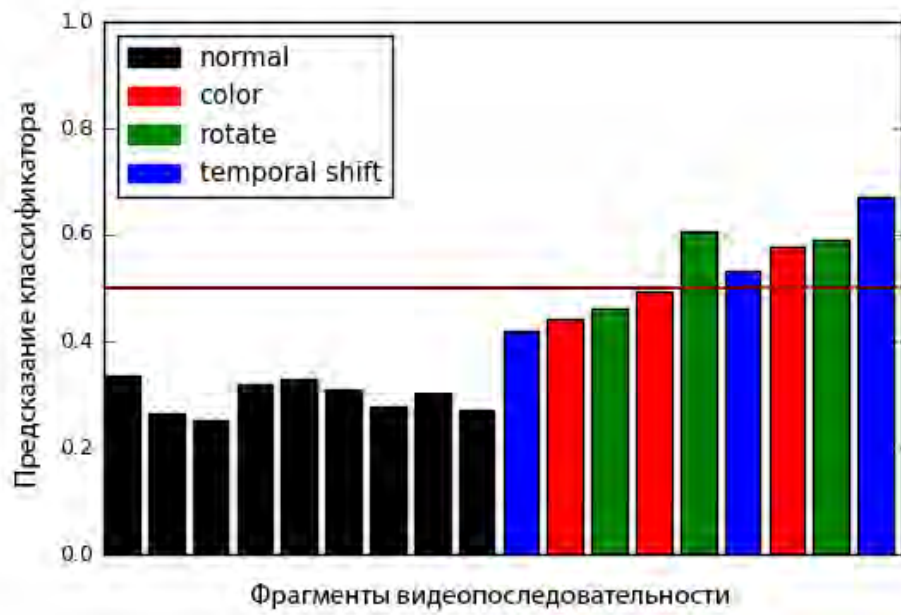
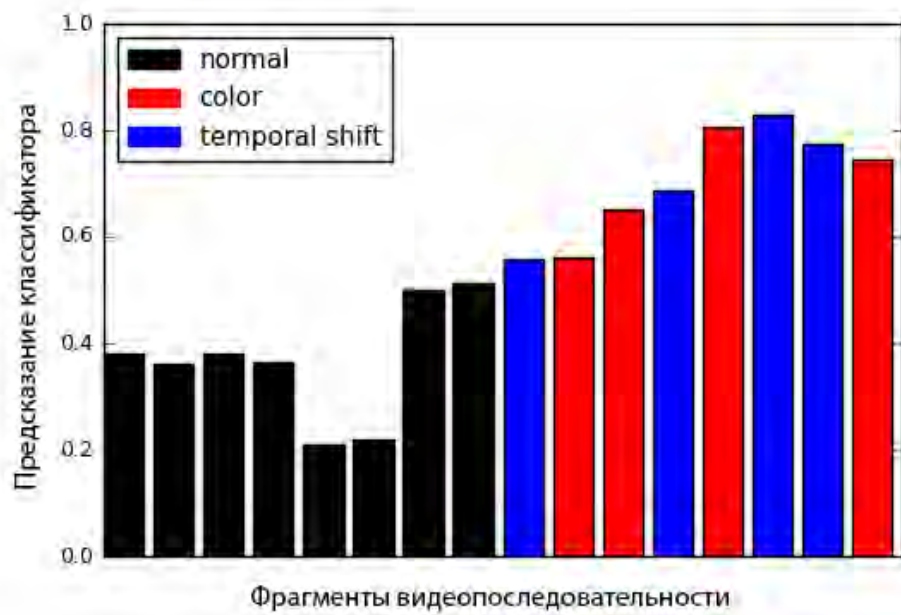


Рис. 16. Обучение на повороте, полный частотный диапазон



$$recall = \frac{TP}{TP + FN}$$

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$F = 2 \frac{precision \cdot recall}{precision + recall}$$

Рис. 17. Обучение на цветовом искажении, β -ритмы

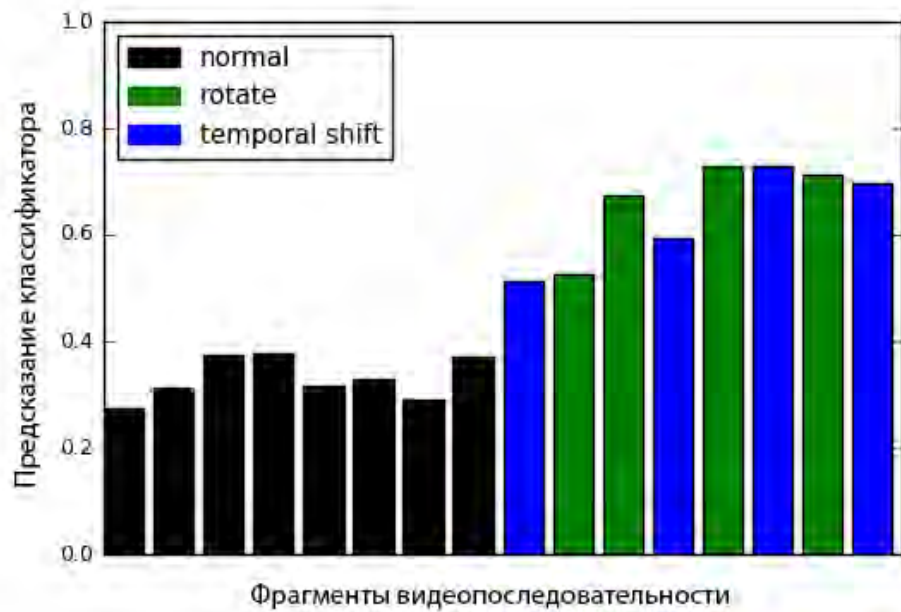
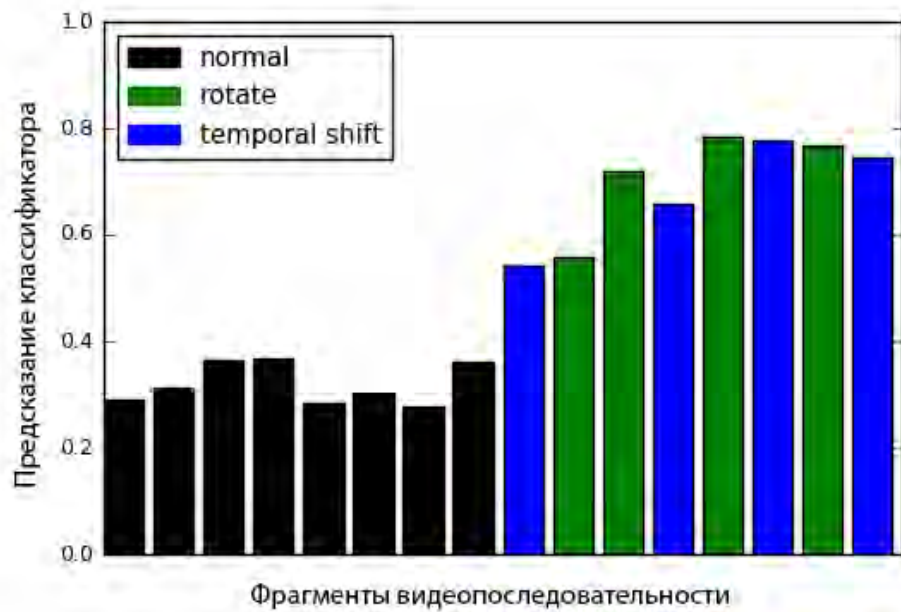


Рис. 18. Обучение на цветовом искажении, полный частотный диапазон



TP — правильно классифицированные объекты первого класса.

TN — правильно классифицированные объекты нулевого класса.

FP — неправильно классифицированные объекты нулевого класса.

FN — неправильно классифицированные объекты первого класса.

Рис. 19. Обучение на повороте, β -ритмы

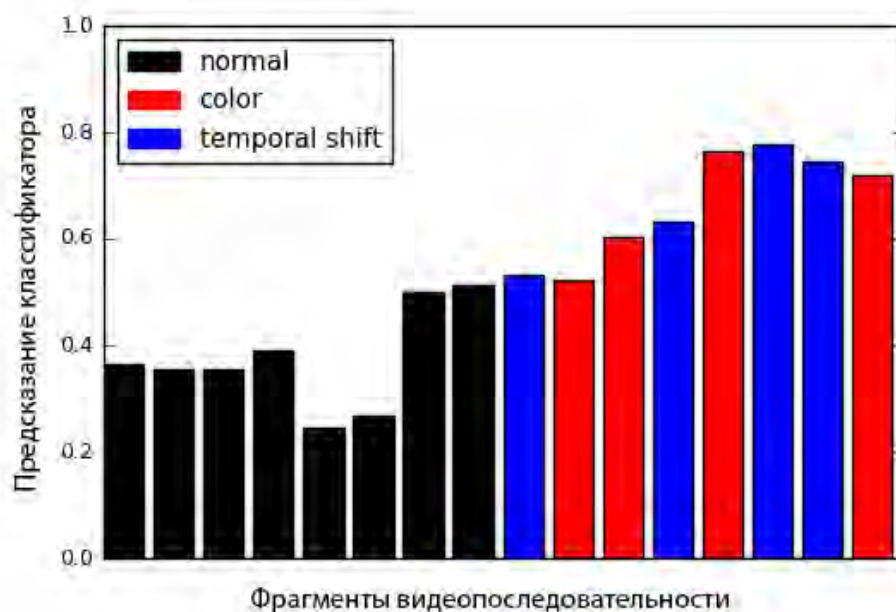
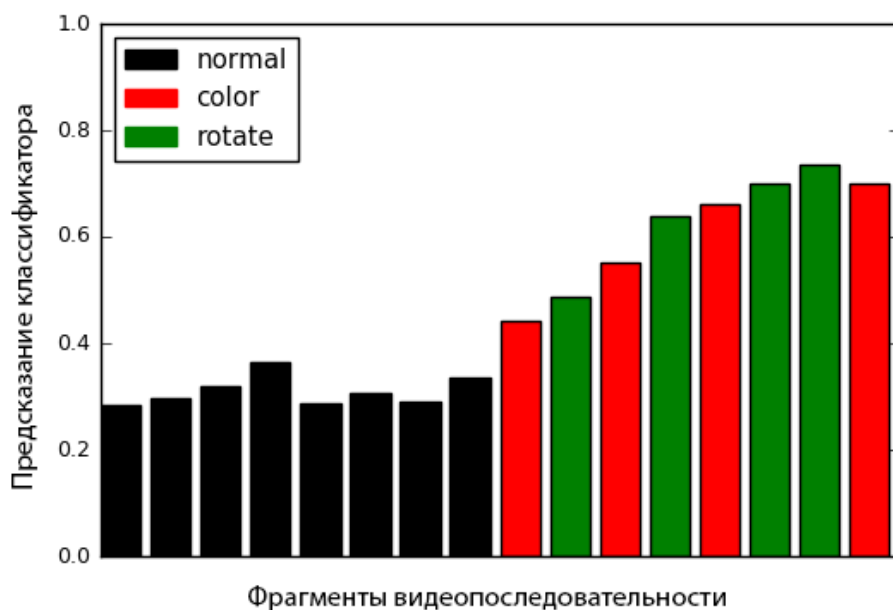


Рис. 20. Обучение на временном сдвиге, β -ритмы



Нулевому классу соответствует видео без ошибок, а первому классу — видео с ошибками. При этом классы и при обучении, и при тестировании являются сбалансированными.

Диапазон α , θ , β ритмов использован как базовый, так как рассмотрение частот, превышающих 38Гц чувствительно к мышечной активности и не является репрезентативным.

Таблица 1. Градиентный бустинг, 500 estimators, результаты для всех интенсивностей ошибок

Частоты	accuracy	precision	F
8 – 14Hz	0.571	0.581	0.564
6 – 38Hz	0.646	0.695	0.605
14 – 38	0.622	0.662	0.577

Таблица 2. Случайный лес, 500 estimators, результаты для всех интенсивностей ошибок

Частоты	accuracy	precision	F
6 – 38Hz	0.638	0.684	0.607
14 – 38Hz	0.612	0.645	0.577

Accuracy необходимо в приведённых таблицах для того, чтобы удостовериться в сбалансированности ответов и ошибок классификатора по классам. В то время, как *precision* отображает точность, по которой и будет оцениваться значимость артефактов 3D при влиянии на зрителя.

В изученной литературе по теме проведённого исследования не было найдено достаточно точного анализа зависимости усталости от ритмов. Было выявлено, что для α -ритмов, определённых для частот от 8 до 14Гц, не прослеживается выделение закономерностей. Ответы классификатора соизмеримы со случайными ответами. То есть на данном диапазоне исследовать усталость не представляется необходимым. Отметим также что α -ритмы соответствуют спокойному состоянию бодрствования.

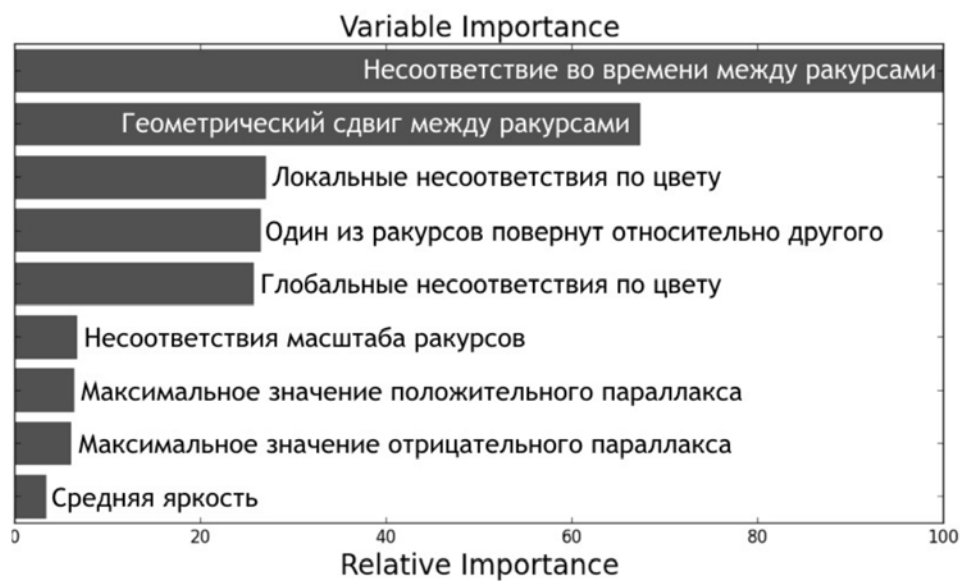
Результаты, приведенные в таблице, трактуются следующим образом: в тестовой выборке содержались ошибки различной интенсивности и фрагменты видео без ошибок. Из графиков видно, что наименьшая интенсивность искажений не была отнесена классификатором к классу ошибок, то есть данной интенсивности недостаточно, для ощутимого влияния на самочувствие человека.

Для β -ритмов характерно состояние бодрствования, повышенное внимание, воздействие внешними стимулами. Сравнивая результаты по α и β -ритмам, можно сделать вывод, что более репрезентативными являются β -ритмы.

Что же касается ранжирования значимости ошибок по их типам, можно сделать вывод,

что наименее значимой ошибкой является цветовое искажение. Сравнение ошибок, связанных с временным сдвигом и поворотом, зависит от интенсивности искажения (при различных интенсивностях значимость одной ошибки может превосходить значимость другой), но в целом можно считать их приблизительно одинаковыми. Данные выводы коррелируют с ранее проведенными экспериментами методом анкетирования [20].

Рис. 21. Относительная степень влияния артефактов и параметров стереовидео на самочувствие зрителя. Результаты анкетирования из работы [20]



7. Заключение

- Было предложено и реализовано решение выделения некачественных фрагментов видео по данным ЭЭГ.
- Были проведены эксперименты и собран набор данных из 17 испытуемых о влиянии видео на самочувствие зрителя.
- Проведён анализ и обработка полученных данных. Выявлено, как можно убрать шумовые данные, обусловленные как звуковой составляющей эксперимента, так помехами при моргании и движении глаз.
- Проанализировано, какие частоты являются наиболее репрезентативными при решении данной задачи.

- Получены результаты альтернативные субъективной оценке при анкетировании о влиянии ошибок на самочувствие.

О результатах данной работы было доложено на конференции «Ломоносов-2017». Тезис работы опубликован в сборнике тезисов конференции [22].

8. Список литературы

- [1] *Сахаров В.Л., Андреевко А.С.* Методы математической обработки электроэнцефалограмм — Учебное пособие. - Таганрог : «Антон», 2000.
- [2] *Марпл – мл. С.Л.* Цифровой спектральный анализ и его приложения — Пер. с англ.- М.: Мир, 1990.
- [3] *Сергиенко. А.Б.* Цифровая обработка сигналов — СПб.: Питер, 2002.
- [4] *Дьяконов В.* MATLAB. Обработка сигналов и изображений. Специальный справочник. — СПб.: Питер, 2002.
- [5] Семинар Центра «Нейроэкономики и когнитивных исследований» Технология обработки сигнала в ЭЭГ реального времени в парадигме нейрообратной связи.
- [6] Электронный ресурс <http://portal.tpu.ru:7777/SHARED/v/VOS/study/disc1/Tab/tema09.pdf>
- [7] Электронный ресурс A Review of Brain-Computer Interface Games and an Opinion Survey from Researchers, Developers and Users
- [8] Электронный ресурс Brain-computer Interfaces: Military, Neurosurgical, and Ethical Perspective
- [9] *А.И. Солонина, Д.А. Улахович, С.М. Арбузов, Е.Б. Соловьева.* Основы цифровой обработки сигналов: курс лекций / – 2-е изд., испр. и перераб. – СПб.: БХВ-Петербург, 2005. – 768 с.
- [10] *А.В. Коломейцева, Г.В. Мишугова, А.П. Мул, Г.Ю. Рябык.* Применеме вейвлет-преобразования и метода Прони для идентификации биогенных сигналов — Вестник ДГТУ, 2010. Т. 10. №4(47)
- [11] *В.В. Шульговский.* Основы нейрофизиологии Аспект Пресс, 2000. - с. 277.
- [12] *В.В. Гнездецкий.* Обратная задача ЭЭГ и клиническая электроэнцефалография Москва, «МЕДпресс-информ», 2004.
- [13] *Г. Шеперд.* Нейро-биология Москва, «Мир», 1987.

- [14] *R. Karlsson, M. Eriksson.* Brain Pattern Recognition: An evaluation of how the choice of training data affect classification accuracy for inexperienced BCI-users, 2014
- [15] *R. Horlings, D. Dactu, LJM Rothkrantz.* Emotion recognition using brain activity ACM, 2008.
- [16] *Н.Н. Красильников, О.И.Красильникова* Исследование критического значения экранного параллакса Information & Control Systems, 2016.
- [17] *Страница проекта VQMT3D: <http://compression.ru/video/vqmt3d/>*
- [18] *Kazuhiko Ukai and Peter A. Howarth.* Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations // In Displays 29, 2008, P. 106–116.
- [19] *Series BT.* Principles for the comfortable viewing of stereoscopic three-dimensional television (3dtv) images 2014.
- [20] *А.В. Анциферова.* Метод определения степени усталости, вызываемой просмотром стереовидео конференция «Ломоносов-2017», секция «Вычислительная математика и кибернетика»
- [21] *Yijun Wang, Shangkai Gao, Xiaorong Gao.* Common Spatial Pattern Method for Channel Selection in Motor Imagery Based Brain-computer Interface Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 2005.
- [22] *С.В. Илларионова.* Нейрокомпьютерный интерфейс в задаче распознавания ошибок 3D-видео конференция «Ломоносов-2017», секция «Вычислительная математика и кибернетика»