



Московский государственный университет имени М.В. Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра Автоматизации Систем Вычислительных Комплексов

Анциферова Анастасия Всеволодовна

**Автоматический метод оценки степени
усталости от просмотра 3D-видео**

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель:
к.ф.-м.н., старший научный сотрудник
Д.С. Ватолин

Москва, 2016

Автоматический метод оценки степени усталости от просмотра 3D-видео

Анциферова Анастасия

Главная цель данной работы заключается в разработке системы, которая будет производить оценку степени дискомфорта, возникающего при просмотре стереофильма зрителем. Система основана на анализе технических параметров сцен стереофильма и использует их для оценки уровня дискомфорта зрителей. В работе проведено исследование влияния различных видов искажений стереоскопического видео на самочувствие человека. На этапе исследования проведено измерение мозговой активности с помощью электроэнцефалограммы, что позволило получить объективную информацию о реакции человека на просмотр стереофильма. Результатом работы является программа, которая на основе данных о качестве стереоскопического видео дает оценку степени усталости зрителя от просмотра анализируемого стереофильма.

Automatic method for fatigue estimation after watching 3D (Three-Dimensional)-video

Antsiferova Anastasiia

The main goal of this investigation is to develop a system that will evaluate the degree of discomfort that appears while watching stereoscopic video. The system is based on an analysis of the technical parameters of stereoscopic scenes and it uses them to assess the discomfort level of the audience. The paper studied the effect of different types of stereoscopic video distortions on human health. On the stage of the study measuring the brain activity was conducted using EEG technology, which made it possible to obtain objective information about the human response to watching stereoscopic video. The result of this work is a computer program that is based on stereoscopic video quality data and estimates the degree of viewer's fatigue which will be obtained during watching the stereoscopic video.

Содержание

1	Введение	4
2	Постановка задачи	7
3	Обзор существующих методов	9
4	Предложенный метод	11
4.1	Подготовка и проведение экспериментов	11
4.1.1	Подготовка стереовидео для экспериментов	12
4.1.2	Проведение экспериментов для получения субъективных данных	15
4.1.3	Проведение экспериментов для получения объективных данных	15
4.2	Обработка и интерпретация результатов	17
4.3	Построение автоматической системы	26
4.4	Результаты	27
5	Описание практической части	31
5.1	Генерация стереовидео для экспериментов	33
5.2	Обработка данных ЭЭГ	33
5.3	Анализ качества стереовидео с помощью программ проекта Video Quality Measurement Tool 3D	34
5.4	Обучение и тестирование искусственной нейронной сети	34
6	Заключение	36
6.1	Результаты работы	36
6.2	Публикации и гранты	36
	Список литературы	37

1 Введение

Технологии стереоскопического кинематографа имели широкое развитие в СССР, однако в наше время рынок 3D находится в стагнации. Многие зрители после просмотра фильма в стереоскопическом формате выходят из кинозала с ощущением головной боли, усталости и других неприятных симптомов, после появления которых стереоскопические фильмы теряют для них привлекательность. Основными причинами плохого самочувствия, возникающего из-за просмотра стереофильмов, являются некачественное оборудование и некачественный видеоконтент. Проблема некачественного оборудования может быть решена выбором другого кинотеатра или техники для просмотра стереофильмов. Однако повышение качества видеоконтента в основном может проводиться непосредственно во время создания стереофильма. Поэтому важным этапом процесса производства стереоскопического фильма является контроль качества данного фильма.



Рис. 1: Две видеокамеры, закрепленные на стереориге

Съемка стереофильма проводится с помощью двух видеокамер (Рис. 1), каждая из которых снимает видео, предназначенное для одного (правого или левого) глаза. Такой способ производства имеет свои преимущества и недостатки: полученное стереоскопическое видео наиболее реалистично для зрительной системы человека, однако какие-либо несоответствия между двумя снятыми видео невозможны для человеческого зрения и мозга. Так, если во время съемки одна из камер повернется относительно другой, такое искажение вероятно будет неестественно и болезненно для зрителя. Исправить по-

вернутое видео можно, но обнаружение такого искажения не столь простая задача.

Допустим, стереограф — человек, контролирующий качество стереофильма во время его производства — посмотрел только что снятый видеофрагмент, на котором несколько кадров оказались искаженными. Во-первых, из-за слабо развитого инструмента автоматического анализа качества стереофильмов, стереограф вынужден полагаться исключительно на свое профессиональное зрение. Во-вторых, помимо искажения, при котором одно из снятых видео оказалось повернутым относительно другого, существует множество других несоответствий в стереоскопических видео: несоответствие по масштабу (когда одно из видео «увеличено» относительно другого), несоответствие по цвету (использование разных светофильтров, блики и тени, попавшие только в одну видеокамеру), рассинхронизация камер (когда при одновременном включении одна из камер дает большую или меньшую задержку), и многие другие [7]. Именно поэтому если стереограф заметит искажение в этой сцене, для его исправления необходимо будет определить его тип. И наконец, в большинстве фильмов такие искажения могут встречаться лишь в нескольких кадрах, и потому могут остаться неисправленными. Стоит также упомянуть, что восприятие стереоизображений у всех людей индивидуально, о чем говорят результаты социальных исследований и опросов [1]. Это значит, что один и тот же фрагмент стереовидео может вызывать различную реакцию зрительной системы у разных зрителей. Поэтому описанная выше субъективная оценка качества стереофильмов, имеющая широкое распространение в наши дни, не позволяет проводить необходимый контроль качества стереоскопических фильмов.

Использование же имеющихся автоматических средств контроля качества стереофильмов нуждается в уточнении. Вернемся к ситуации, когда необходимо проанализировать только что снятый фрагмент стереофильма на наличие искажений. Допустим, используемая автоматическая система нашла множество различных несоответствий, нуждающихся в исправлении. Однако, как это часто бывает, бюджет и сроки производства стереофильма не позволяют исправить все несоответствия, а потому необходимо принять решение, какие из обнаруженных искажений необходимо исправить в первую

очередь. Итак, мы наконец подошли к главной цели данной работы: необходимо научиться определять самые болезненные для зрителей искажения в стереофильмах. Таким образом, при исправлении самых заметных и болезненных для человека несоответствий в стереоскопических фильмах уменьшится число зрителей, испытывающих головную боль и дискомфорт во время просмотра стереофильма, даже если в этом фильме не все несоответствия будут исправлены.

2 Постановка задачи

Главной целью данного исследования является разработка метода определения уровня дискомфорта, который испытает зритель при просмотре стереофильма. Система должна давать оценку усталости автоматически путем анализа особенностей стереофильма.

В данной работе будут использованы следующие понятия:

- Стереofilm, стереовидео — видео в стереоскопическом формате, то есть два видео, каждое из которых предназначено для одного глаза.
- Стереопара — кадр стереоскопического видео в определенный момент времени, то есть один кадр из видео, предназначенного для левого глаза, и один — для правого.
- Ракурс (левый или правый) — один из кадров стереопары или одна из двух видеопоследовательностей стереоскопического фильма.
- Искажение, несоответствие между ракурсами стереовидео — в данной работе под этими терминами подразумевается любое несоответствие между видеорядами или кадрами стереофильма, которое не имеет отношения к естественному горизонтальному смещению видеокамер, обеспечивающему стереоэффект.
- Усталость, дискомфорт, головная боль — в данной работе эти термины обозначают негативное самочувствие зрителя во время просмотра стереофильма.
- Параллакс — изменение видимого положения объекта относительно удаленного фона в зависимости от положения наблюдателя.

Автоматическая система определения уровня усталости зрителя, вызванной просмотром стереофильма, должна по данным качества стереофильма давать оценку степени дискомфорта, который будет испытывать зритель при просмотре этого стереофильма. Для получения необходимых данных качества стереовидео будут использоваться автоматические средства оценки качества стереофильмов, разработанные в рамках проекта VQMT3D [18] лабо-

ратории «Компьютерной графики и мультимедиа» МГУ. Используемые средства позволяют измерить величину несоответствий между ракурсами стереовидео, которые возникают в процессе съемки стереофильма. В данной работе будет проводиться исследование влияния на усталость четырех следующих видов несоответствий: расхождение ракурсов по цвету, масштабу, поворот одного из ракурсов относительно другого и несоответствие по времени. Для учета отдельных особенностей кадров стереофильма при анализе система также будет использовать данные о величине параллакса и яркости кадра.

Для получения данных о реакции зрителя на просмотр стереовидео, содержащего различные несоответствия в стереопаре, необходимо будет провести серию экспериментов. В экспериментах будет использоваться специально подготовленное стереовидео, содержащее различные виды и различную степень расхождения между ракурсами. На этапе исследования будет использовано измерение мозговой активности с помощью электроэнцефалографа, что позволит получить объективную информацию о физической реакции человека на просмотр стереофильма. Продукт будет представляться в виде программы. По полученным данным качества стереофильма программа будет выдавать для каждого кадра число — уровень усталости (по шкале от «1» до «10», где «10» — нет усталости, «1» — сильная усталость), которую испытает среднестатистический зритель после просмотра проанализированного видео. Система будет производить диагностику каждого кадра стереофильма на наличие болезненных для человека несоответствий между ракурсами стереофильма. Система будет востребована стереографами, как автоматическое средство контроля качества стереофильма, учитывающее особенности восприятия ошибок стереоформата зрительной системой человека, так же система может найти применение в других областях, использующих стереотехнологии.

3 Обзор существующих методов

Рост доступности и развития стереотехнологий способствовали широкому изучению данной предметной области. Текущая работа начиналась с анализа диссертации французского ученого Wei Chen [3], подробно описавшего особенности стереовидео и его влияния на людей. Автор провел субъективное сравнение дискомфорта зрителя от различных искажений стереоскопического видео, а так же использовал технологию ЭЭГ для сравнения восприятия зрителем двумерного и трехмерного видео. В этой работе выявлены закономерности зрительного восприятия геометрических и других несоответствий между ракурсами стереовидео, однако не проводилось относительное сравнение влияния на зрителя различных видов несоответствий. Использование технологии ЭЭГ для сравнения реакции человека на просмотр видео в форматах 2D и 3D проводили и корейские ученые [12], которые отметили увеличение мозговой активности (в особенности бета-ритма) при просмотре объемного видео по сравнению с двумерным. Схожие результаты были получены в работе [8], где помимо анализа ритмов оценивались их определенные комбинации. В данных работах оценивается только общее состояние зрителя до и после просмотра стереовидео и не исследуется зависимость появления усталости от особенностей стереофильма.

При анализе электроэнцефалограммы с целью выявления в ней признаков усталости допустимо абстрагироваться от типа задачи, выполняемой участниками во время эксперимента. В работах [13],[9],[11] авторами предложены алгоритмы автоматического определения уровня усталости по данным ЭЭГ, в которых используются различные методы машинного обучения и характеристики сигнала. При обучении предложенных моделей использовались данные электроэнцефалограммы, записанные при выполнении различных когнитивных задач. Более глубокое изучение причин появления усталости, возникающей именно при просмотре стереофильмов, требует детального анализа особенностей стереовидео, и существует ряд исследований зависимости появления усталости от определенных параметров сцен стереофильмов. Так, в работе [19] рассматривается влияние движения в салиентных областях сцены на уровень комфорта ее визуального восприятия. Авторы рассмотрели все

виды движения объектов в сцене: по вертикали, горизонтали и движение по глубине, то есть изменение видимого расстояния до объекта, обеспечиваемое стереоэффектом. Последний из этих видов движения был изучен в работе [4], авторы которой помимо субъективного оценивания уровня дискомфорта измеряли частоту моргания зрителя как объективный показатель усталости глаз. Построенные в ходе исследований модели не учитывают движения камеры в сцене, однако с их помощью было выявлено отрицательное влияние высокой скорости движения объектов по глубине на дискомфорт зрителей.

Несмотря на большое количество исследований в изучаемой области, на данный момент не существует автоматических алгоритмов оценки усталости, возникающей из-за несоответствий между ракурсами стереофильма. Безусловно, особенности сцены могут влиять на самочувствие зрителя, однако в первую очередь необходимо избежать невозможных для человеческой зрительной системы искажений. В то же время объективная оценка усталости, вызванной просмотром стереофильма с такими искажениями, должна быть использована для автоматического предсказания степени дискомфорта, который испытает зритель при просмотре.

4 Предложенный метод

Решение поставленной задачи разделяется на два основных этапа: проведение экспериментов для получения данных об усталости зрителя при просмотре стереовидео и разработка программного средства, с помощью которого такую усталость можно будет предсказывать. Схема проведенного исследования изображена на Рис. 2. Далее будет приведено подробное описание подготовки и проведения экспериментов.

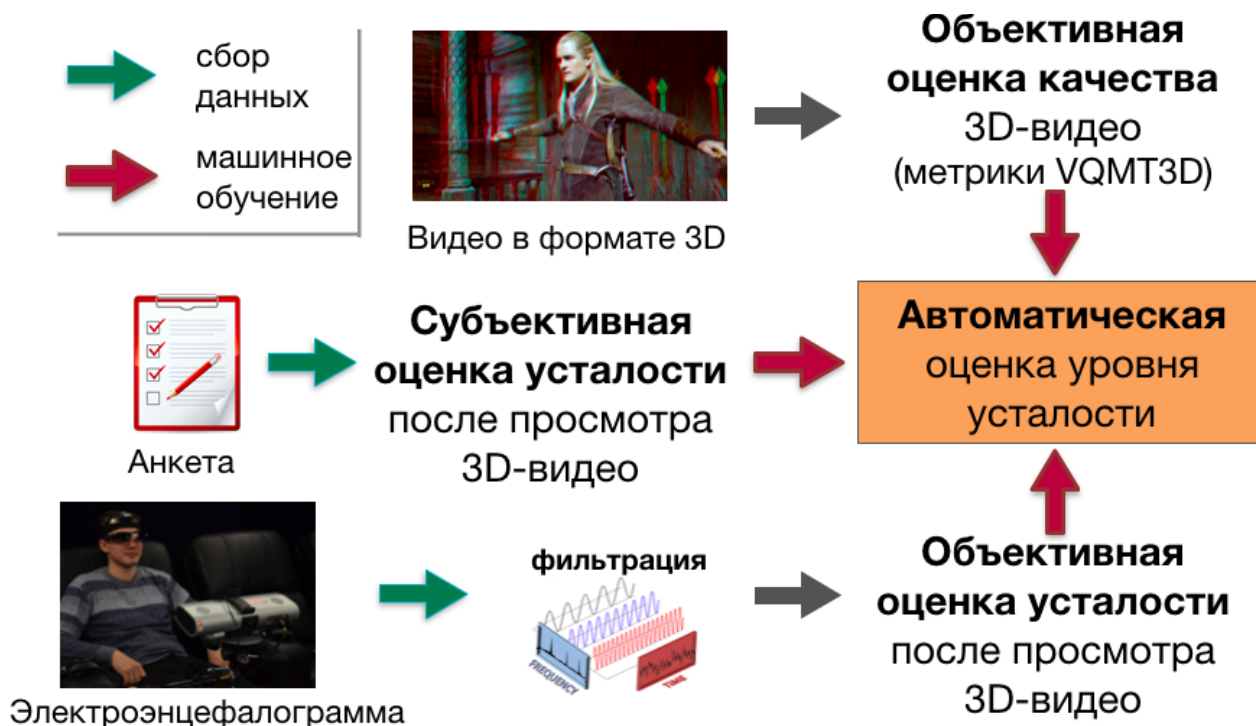


Рис. 2: Общая схема метода

4.1 Подготовка и проведение экспериментов

Для получения информации об усталости и дискомфорте зрителя при просмотре 3D-видео было подготовлено 2 типа экспериментов: для получения субъективной и объективной оценки усталости. Под субъективной оценкой усталости подразумевается собственное ощущение и оценка своего состояния участниками эксперимента. Объективная оценка усталости — это оценка состояния человека по независимым от субъективного восприятия данным. В данном исследовании в качестве объективной оценки использовались данные электроэнцефалограммы (ЭЭГ), записанные во время проведения экспери-

мента. В экспериментах приняли участие 21 человек, 17 из них — в эксперименте по оценке субъективной усталости, 8 — для получения данных ЭЭГ. Среди них 6 женщин, 15 мужчин. Средний возраст — 20 лет. Длительность каждого эксперимента составляла 60 минут. Всего было проведено 11 экспериментов.

4.1.1 Подготовка стереовидео для экспериментов

Для проведения экспериментальной оценки дискомфорта при просмотре стереовидео, содержащего характерные несоответствия между ракурсами, было подготовлено 76 фрагментов из художественных фильмов длительностью 30 секунд каждый. Все выбранные фильмы были произведены методом стереосъемки, то есть каждый из ракурсов снимался отдельной камерой. Такой метод производства стереоконтента является наиболее естественным для зрительной системы человека. Для того, чтобы промоделировать характерные для съемки искажения в стереопаре, к подготовленным видеофрагментам искусственным образом были добавлены:

- несоответствия по цвету между ракурсами стереовидео
- несоответствия по масштабу между ракурсами стереовидео
- поворот ракурсов стереовидео
- несоответствие по времени между ракурсами стереовидео

Обычно при съемке фильма указанные несоответствия появляются с разной степенью. Например, степень поворота одного из ракурсов относительно другого можно описать величиной угла, на который этот ракурс повернут.

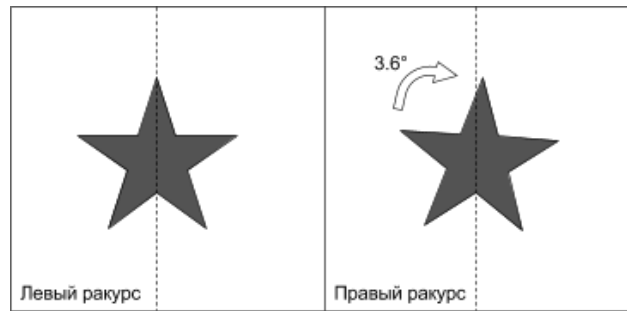
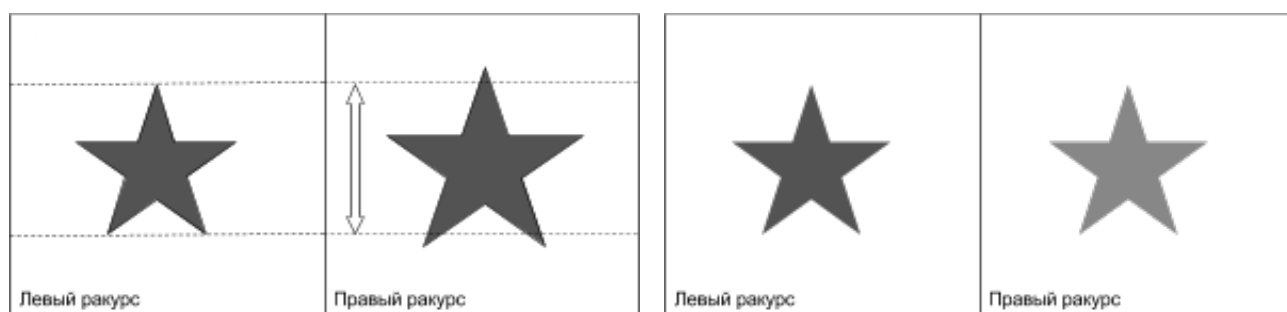


Рис. 3: Схематическая визуализация поворота одного из ракурсов относительно другого.

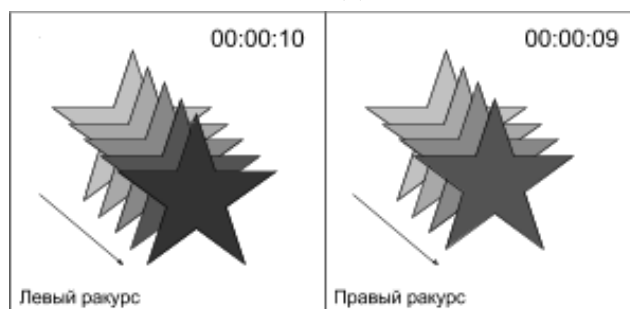
На Рис. 3 схематически показан угол поворота. При этом в различных условиях съемки камеры могут быть повернуты на различные углы. То же самое касается других несоответствий, схематически представленных на Рис. 4: величина различия ракурсов по масштабу может быть разной в зависимости от оптических особенностей оборудования, несоответствия по цвету зависят от освещения сцены и световых фильтров камер, временной сдвиг между ракурсами — от синхронизации камер и иногда даже от используемых элементов питания. Именно поэтому при проведении эксперимента необходимо учесть тот факт, что степень несоответствия между ракурсами может быть разной. Для этого каждое из указанных несоответствий добавлялось к фрагментам экспериментального стереовидео с разной интенсивностью. Условно интенсивность каждого типа несоответствия варьировалась от «0» до «4», где «0» — полное отсутствие искажения стереопары, «4» — сильное искажение стереопары, которое соизмеримо с худшими примерами, встречаемыми в художественных стереофильмах. Значения для разных типов искажений и их градации выбраны в соответствии с указанными в [16]. Для того, чтобы сравнение дискомфорта от разных «уровней» несоответствия было наиболее объективным, каждый тип несоответствия добавлялся к одинаковым фрагментам стереофильма, то есть каждый выбранный фрагмент из фильма присутствовал в тестовом стереовидео 5 раз (со степенью несоответствия от «0» до «4»).

Порядок следования сгенерированных фрагментов стереофильмов был случайным с учетом одинаковых фрагментов стереофильмов, то есть в итоговом стереовидео для эксперимента не было подряд идущих одинаковых



(a) несоответствие масштаба ракурсов стереовидео

(b) несоответствие цвета ракурсов стереовидео



(c) временной сдвиг между ракурсами стереовидео

Рис. 4: Схематическая визуализация несоответствий между ракурсами стереопары.

фрагментов фильмов. Общую схему расположения видео можно условно закодировать следующим образом: буквами A, B, C, D обозначим выбранные фрагменты из художественных фильмов; буквами c, s, r, t , где c (color) — цветовое искажение, s (scale) — несоответствие масштаба, r (rotate) — поворот одного из ракурсов, t (temporal shift) — временной сдвиг между ракурсами, обозначим типы несоответствий в стереопаре, которые искусственно добавляются к фрагментам стереофильмов с различной силой, обозначенной цифрами «0», «1», «2», «3», «4», где «0» — полное отсутствие искажения стереопары, «4» — сильное искажение стереопары. Тогда пример фрагмента схемы стереовидеопоследовательности будет выглядеть так:

$$A1s, D3t, B2c, D4r, C2t, A0c, \dots \quad (1)$$

Итоговое стереовидео было условно поделено на 4 блока (по 19 видеофрагментов): после каждого блока была добавлена 1 минута видео с номером прошедшего блока для того, чтобы участники эксперимента имели возможность

отдохнуть. Также после каждого фрагмента итогового стереовидео было добавлено 9 секунд видео, содержащего черный экран и номер предыдущего видеофрагмента в формате «№ блока.№ фрагмента», где «№ блока» — от 1 до 4, «№ фрагмента» — от 1 до 19. Таких фрагменты были добавлены для того, чтобы в это время участники эксперимента заполнили анкету с вопросами о своем самочувствии. Об анкетах будет подробно рассказано в 4.1.2.

4.1.2 Проведение экспериментов для получения субъективных данных

Как уже упоминалось выше, в данной работе субъективные данные — это данные об усталости зрителя, полученные с помощью анкетирования. Анкетирование проводилось во время просмотра подготовленного стереовидео, описанного в 4.1.1, в специально отведенное для этого время.

Эксперимент проводился в помещении, оборудованном проектором с активной технологией 3D. Разрешение проектора 1200×800 , частота 120 Hz. Высота экрана 1.5 метра, расстояние от места просмотра до экрана 5 метров. Просмотр осуществлялся без освещения, для ответа на вопросы анкеты использовалась электронная форма.

Анкета состояла из одного вопроса, для ответа на который участнику нужно было оценить свое самочувствия по шкале от «1» до «9», где «9» — улучшение самочувствия, «7» — нет дискомфорта, «3» — сильный дискомфорт, «1» — нестерпимо сильный дискомфорт и усталость. На этот вопрос участники должны были ответить после просмотра каждого фрагмента тестового стереовидео в специально отведенное для этого время.

4.1.3 Проведение экспериментов для получения объективных данных

Для получения субъективных данных об уровне усталости зрителя при просмотре стереовидео в данном исследовании использовалась технология электроэнцефалографии. Электроэнцефалография эффективно используется для изучения электрической активности мозга, так как получаемая электроэнцефалограмма имеет высокое временное разрешение и содержит информацию

о реакции головного мозга на внешние раздражители.



(а) Участник отвечает на вопросы анкеты (б) Участник ожидает стабилизации электродов электроэнцефалографа

Рис. 5: Проведение эксперимента для получения данных ЭЭГ.

Эксперимент проводился в специально оборудованном помещении, предоставленном компаниями Movie Research Company [14] и Neurotrend [15]. Помещение оборудовано проектором с активной технологией 3D, а так же 5 мягкими креслами с закрепленными на них полиграфом, трекером глаз и видеокамерой. Примерное расстояние от места просмотра до экрана составляло 8 метров, высота экрана составляла 2 метра. Фотографии помещения и участников, сделанные перед проведением эксперимента, представлены на Рис. 5.

Для записи ЭЭГ использовались беспроводные электроэнцефалографы B-Alert X24 с частотой записи сигнала 1024 Hz. Для записи электрической активности к определенным точкам на голове участника эксперимента крепятся электроды. В проведенном эксперименте электроды крепились по схеме, изображенной на Рис. 6; всего было подключено 15 электродов: Fp1, F7, F8, T4, T3, Fp2, O1, P3, F3, F4, C4, P4, POz, C3, O2.



Рис. 6: Схема расположения электродов по системе 10-20.

Запись ЭЭГ проводилась все время просмотра подготовленного стереовидео. Использование электронных форм для оценки уровня усталости зрителей наряду с записью ЭЭГ невозможно, так как сигналы электронных устройств являются причиной помех при записи сигнала. Поэтому в специально отведенное время между блоками видеопоследовательностей участники отвечали на вопросы об их самочувствии, всего таких опросов было 5: до начала просмотра и после каждого из блоков по 19 фрагментов стереовидео.

4.2 Обработка и интерпретация результатов

Для выявления признаков усталости в полученных данных необходимо определить количественные характеристики, которые будут отражать уровень дискомфорта зрителя. В случае с субъективной оценкой такой характеристикой является непосредственно индекс усталости (число от «1» до «9»), с помощью которого участники эксперимента оценили свое самочувствие. В отличие от анкетирования, данные электроэнцефалограммы представляют только характеристику электрической активности мозга и не могут в явном виде быть использованы для оценки уровня усталости. Данный раздел посвящен описанию способа анализа данных электроэнцефалограммы для выделения количественной характеристики усталости, который был разработан в рамках проведенного исследования.

Несмотря на то, что помещение, в котором проводился эксперимент, специально изолировано, на данных ЭЭГ встречаются участки с помехами. Помехи могут возникать, например, из-за движения участника во время эксперимента, так как мышечная активность также регистрируется электрода-

ми. Именно поэтому первым этапом анализа данных электроэнцефалограммы является фильтрация полученного сигнала. На Рис. 7 изображен фрагмент электроэнцефалограммы одного из участников эксперимента, содержащий помехи.

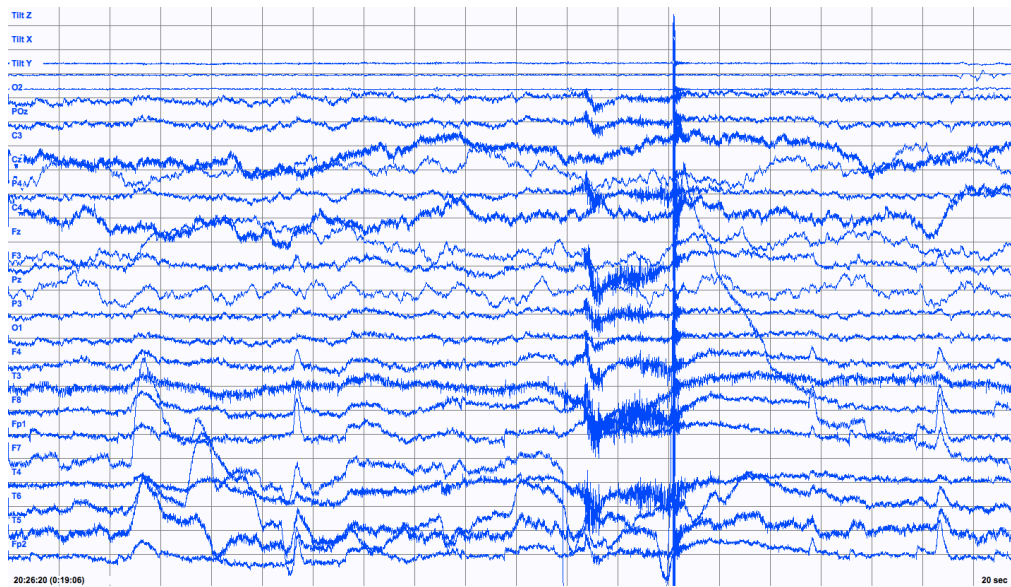


Рис. 7: Электроэнцефалограмма одного из участников эксперимента. По оси X — время, по оси Y — каналы (электроды). Электроэнцефалограмма содержит зашумленный фрагмент, подлежащий удалению.

Так как длительность эксперимента составляла более 60 минут, а запись ЭЭГ велась с частотой 1024 измерения в секунду, просматривать электроэнцефалограммы участников и искать подобные искажения вручную было бы слишком долго. Поэтому для поиска таких испорченных измерений была использована информация из не подключенных электродов, которые работали как приемники внешних сигналов, оказывающих влияние на электроэнцефалограмму, а так же каналы Tilt X, Tilt Y, Tilt Z, содержащие информацию о положении головы участника, полученную с помощью акселерометра, установленного в электроэнцефалографе. Используя «правило трех сигм» для данных перечисленных каналов, была построена маска, используемая в даль-

нейшем анализе ЭЭГ:

$$x_i^j = \begin{cases} 1, & \text{if } x_i \text{ in } [\text{mean}(\text{channel}_j) - 3 * \text{std}(\text{channel}_j); \\ & \text{mean}(\text{channel}_j) + 3 * \text{std}(\text{channel}_j)] \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

Здесь i — номер измерения, j — номер канала ЭЭГ.

Обычно при анализе ЭЭГ рассматривают активность мозга в определенных частотных диапазонах, называемых мозговыми ритмами [8]. Описание этих ритмов будет приведено далее, и первым этапом выделения ритмов является получение частотно-временного представления сигнала ЭЭГ. Для получения информации о частотной активности сигнала ЭЭГ было использовано оконное преобразование Фурье:

$$F(m, \omega) = \sum_{n=-\infty}^{+\infty} f[n] \omega[n - m] * e^{i\omega n} \quad (3)$$

В качестве оконной функции использовалось окно Ханнинга:

$$\omega(n) = 0.5 \left(1 - \cos \left(\frac{2\pi n}{N-1} \right) \right) \quad (4)$$

Здесь f — исходный сигнал, N — длина сигнала, n — номер измерения, m — параметр алгоритма (в данном исследовании использовалось $m = 2$).

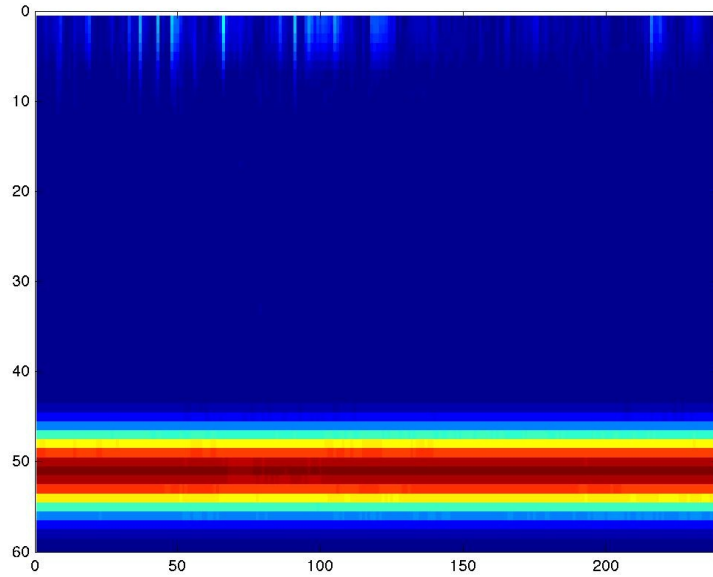


Рис. 8: Спектрограмма сигнала ЭЭГ, на которой видно сильное преобладание частоты 50 Гц. По оси X — время, по оси Y — частота (по убыванию).

В работе был выбран размер окна, равный 1024, и полученная матрица значений содержит спектр частот, равный половине размера окна, согласно теореме Котельникова. Визуализация полученной спектрограммы представлена на Рис. 8.

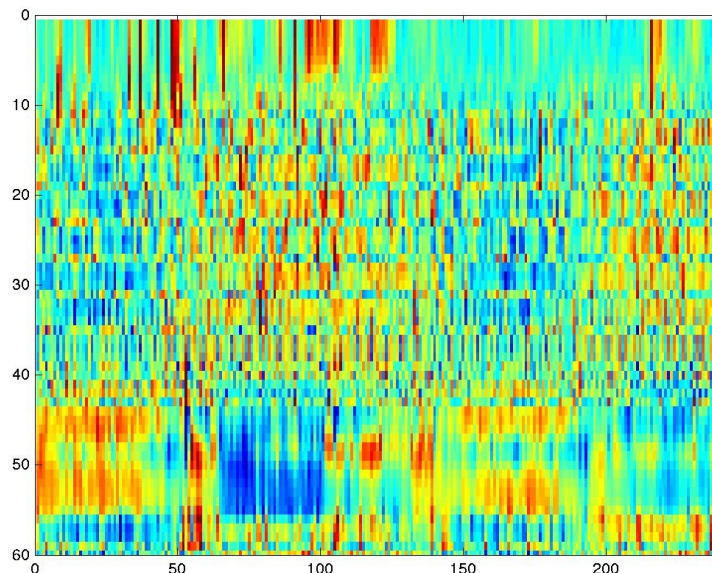


Рис. 9: Спектрограмма сигнала ЭЭГ после фильтрации частоты 50Hz. По оси X — время, по оси Y — частота (по убыванию).

На спектрограмме видно преобладание частоты 50 Hz, которая перекрывает остальные частоты. Это происходит из-за того, что записывающие приборы работают на этой частоте. Обычно при фильтрации ЭЭГ эту частоту обнуляют, но для сохранения информации была проведена фильтрация с применением уже упомянутого «правила 3-х сигм», где среднее значение вычислялось по спектру частот в каждый момент времени, а значения частоты 50 Hz, выходящие за границы посчитанного среднего значения более, чем на 3 значения стандартного отклонения, заменялись соответствующим значением среднего плюс/минус 3 стандартных отклонения. Измененная спектрограмма изображена на Рис. 9.

Теперь на полученной спектрограмме можно наблюдать временную активность в различных частотных диапазонах, или мозговых ритмах. Как упоминалось ранее, существует несколько диапазонов частот, каждый из которых связывают с определенной активностью или состоянием человека:

- 1-4 Hz — дельта-ритм (обычно ассоциируется с фазой глубокого сна)

- 4-8 Hz — тета-ритм
- 8-13 Hz — альфа-ритм
- 13-30 Hz — бета-ритм (обычно ассоциируется с активной умственной деятельностью)
- 30-50 Hz — гамма-ритм

Далее будет приведено описание количественной характеристики, которая будет использоваться для определения величины частотной активности. Однако сначала необходимо обратить внимание на то, что представление сигнала в частотно-временной области дает более широкую картину для анализа ЭЭГ, чем обычное частотное представление. Например, применив обычное преобразование Фурье (или взяв за ширину окна в оконном преобразовании длину всего сигнала) получится частотное представление, показанное на Рис. 10.

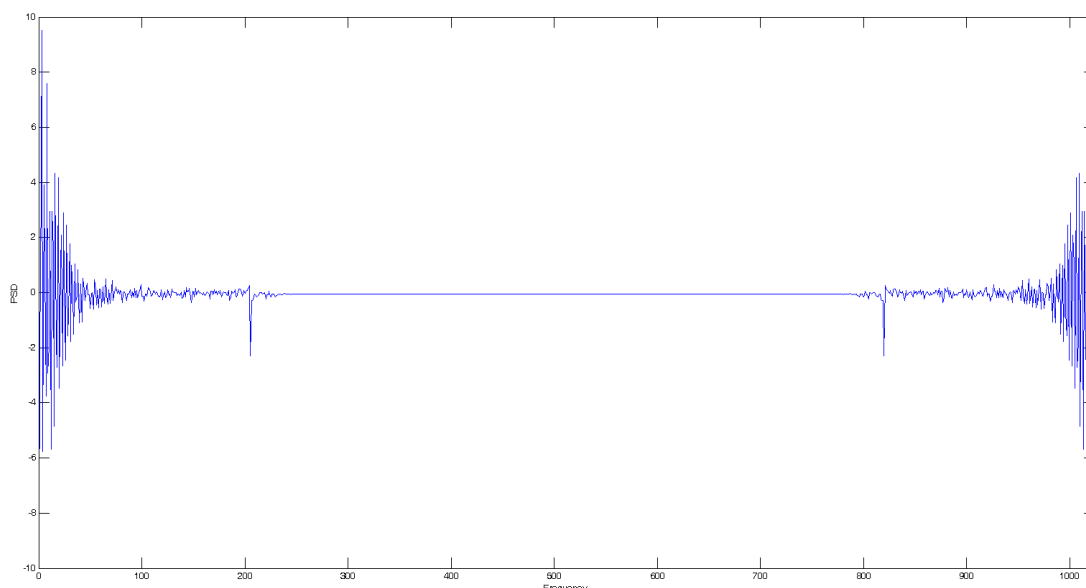


Рис. 10: Частотное представление фрагмента электроэнцефалограммы. По оси X обозначены частоты. График симметричен относительно середины, так как результат преобразования Фурье — комплексная амплитуда синусоидальной функции (по оси Y), поэтому в правой части графика находятся комплексно-сопряженные значения, и по теореме Котельникова анализируется только половина частотного спектра.

Для определения значения, которое будет характеризовать величину «активности» частоты, вычисляется значение спектральной плотности мощности сигнала (Power Spectral Density, PSD):

$$PSD = \frac{|F(k)|^2}{N^2} \quad (5)$$

где $F(k)$ — комплексная амплитуда синусоидального сигнала (результат преобразования Фурье), N — длина сигнала. Вычислив данное значение и приведя спектр к правильной размерности, он принимает вид, изображенный на Рис. 11.

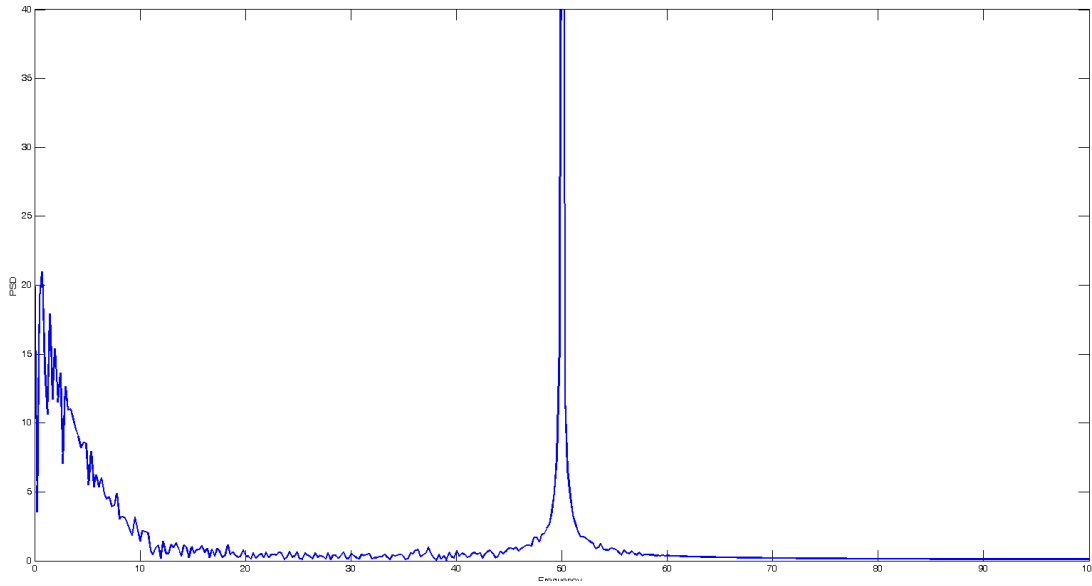
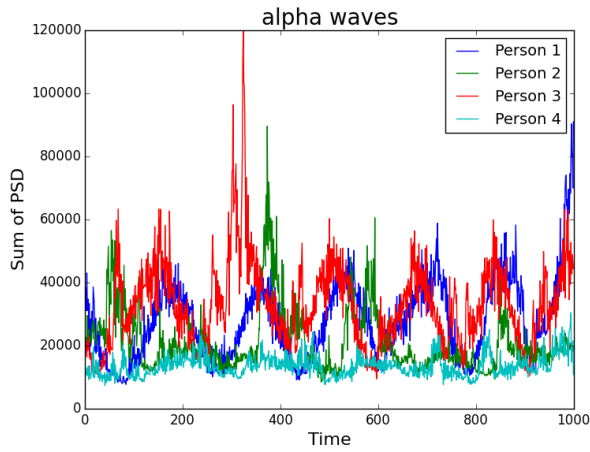


Рис. 11: Спектральная плотность мощности сигнала. По оси X — частота, по оси Y — значение PSD . Здесь виден пик на частоте 50 Hz, на которой работают приборы.

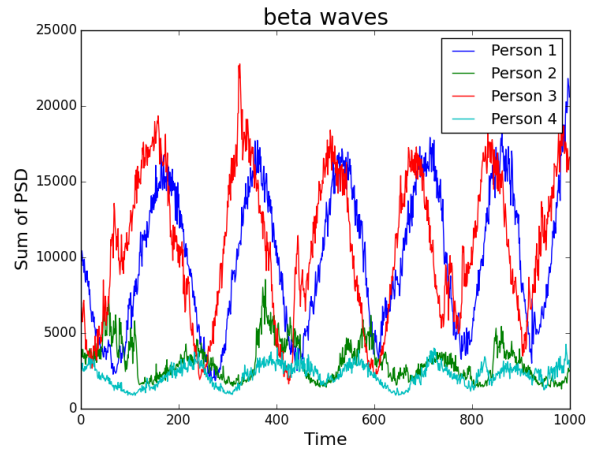
На таком графике так же видна активность ритмов, расположенных в определенных частотах. Для того, чтобы учесть вклад каждой частоты из диапазона, соответствующего ритму, спектральная плотность мощности суммируется в данном частотном диапазоне:

$$Sum_{PSD} = \sum_{k \subseteq \alpha, \beta, \theta, \delta, \gamma - rhythms} PSD(k) \quad (6)$$

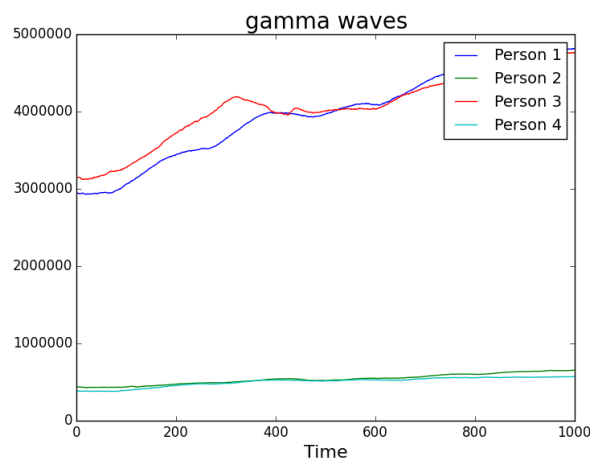
где k — диапазон частот, соответствующий перечисленным мозговым ритмам. Визуализация суммарной спектральной плотности мощности, вычисленной на одном из каналов ЭЭГ, приведена на Рис. 12.



(а) Альфа-ритм



(б) Бета-ритм



(с) Гамма-ритм

Рис. 12: Суммарная спектральная плотность мощности мозговых ритмов в зависимости от времени. Каждая линия соответствует данным одного участника эксперимента.

Полученные значения характеристики ритмов могут быть использованы для оценки реакции зрителя на просмотр стереовидео, однако необходимо учесть еще одну особенность человеческого восприятия. Просмотр неестественного для зрения фрагмента стереофильма может вызвать дискомфорт, который зритель будет испытывать еще некоторое время после непосредственного просмотра искаженного кадра, ровно как и сама реакция на просмотр не мгновенная и варьируется в диапазоне от 60 до 200 миллисекунд (компоненты зрительных вызванных потенциалов [17]). Поэтому следующим этапом анализа ритмов является вычисление их энтропии, то есть меры непред-

сказуемости значения ритма.

В данной работе анализируется энтропия ритмов, вычисленная по алгоритму approximate entropy [6] следующим образом: Исходный сигнал:

$$u_1, \dots, u_n \quad (7)$$

1. Формируются векторы:

$$x_1, \dots, x_{N-m+1} : [x_i = u_i, u_{i+1}, \dots, u_{i+m-1}] \quad (8)$$

2. Обозначается

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log(C_i^m(r)), \quad (9)$$

где

$$C_i^m(r) = [\text{количество } x(j) \text{ таких, что } d[x(i), x(j)] < r] / (N - m + 1) \quad (10)$$

$$d[x, y] = \max |x(a) - y(a)| \quad (11)$$

3. Энтропия равна

$$ApEn = \Phi^m(r) - \Phi^{m+1}(r) \quad (12)$$

Полученное значение — энтропия суммарной спектральной плотности мощности мозговых ритмов - используется в данной работе в качестве объективной характеристики степени усталости зрителя, вызванной просмотром стереофильма. Ее объективность подтверждается Рис. 13, на котором изображено значение данной характеристики до и после просмотра стереовидео, а так же Рис. 14, на котором изображена субъективная оценка зрителями своего самочувствия и объективное значение предложенной характеристики усталости. На этих графиках изображено суммарное значение энтропии, посчитанное на данных 30-секундных фрагментов электроэнцефалограммы до начала просмотра, а так же после просмотра каждого блока тестового стереовидео; во время записи этих фрагментов участники эксперимента сидели неподвижно с закрытыми глазами.

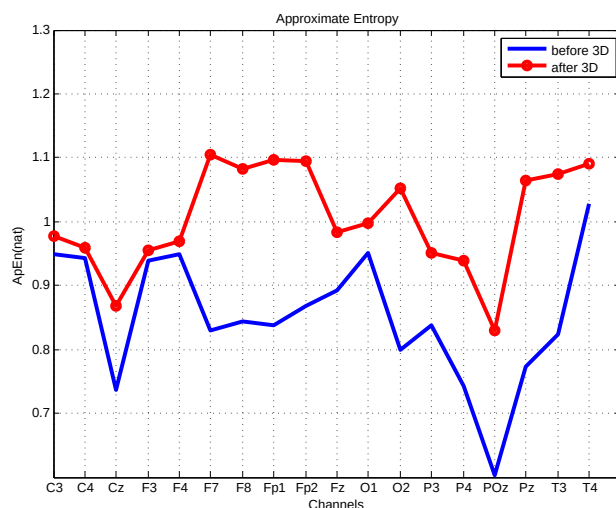


Рис. 13: Значение предложенной характеристики усталости зрителя до и после просмотра зрителем стереовидео. По оси X — различные каналы ЭЭГ, по оси Y — значение энтропии. Красной линией обозначено суммарное значение энтропии после просмотра стереовидео, синей — до просмотра стереовидео.

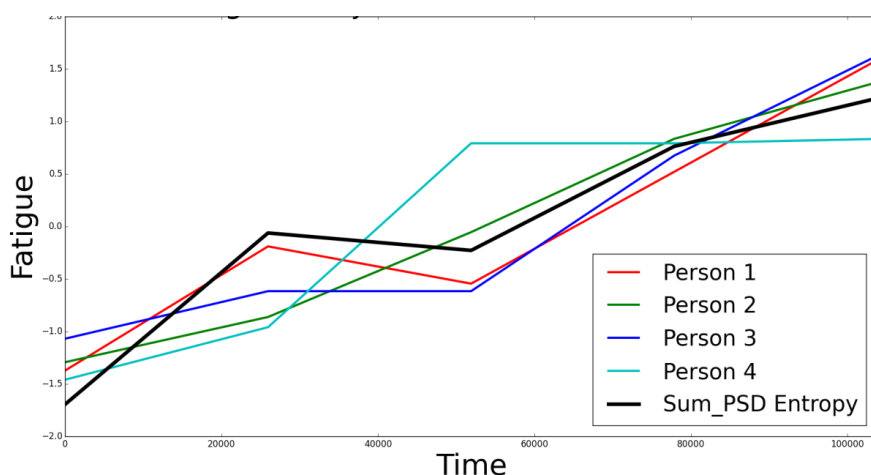


Рис. 14: Сравнение субъективной и предложенной объективной оценки усталости. По оси X — время просмотра стереовидео, по оси Y — степень усталости зрителя. График построен по данным канала №13 - P4.

Энтропия суммарной спектральной плотности мощности была посчитана для каждого кадра видео в окне, равном 4 секундам. Таким образом, длина вектора характеристики активности ритмов совпадает с длиной видео, то есть каждому кадру просмотренного стереовидео можно поставить в соответствие значение энтропии суммарной спектральной плотности мощности ритмов, в то время как субъективная оценка усталости, полученная с помощью анке-

тирования, содержит всего лишь 1 значение на 30 секунд стереовидео.

4.3 Построение автоматической системы

Поставленная задача требует от системы оценку степени усталости, которая является вещественным числом и вычисляется по значениям величины несоответствий между ракурсами стереофильма. Данная задача является задачей регрессии и решается с помощью машинного обучения.

Обучение проводилось отдельно для данных субъективного и объективного экспериментов. Все полученные данные были поделены поровну на две части — обучающую выборку и тестовую выборку. В ходе разработки системы было испробовано несколько методов машинного обучения, среди которых были градиентный бустинг, решающие леса, искусственные нейронные сети. Однако наилучший результат был достигнут при использовании рекуррентной искусственной нейронной сети. Этот факт объясняется временной зависимостью данных, которая учитывается рекуррентной нейронной сетью.

Признаками для нейронной сети являются значения автоматических алгоритмов оценки качества стереовидео [18], пример визуализации которых изображен за Рис. 15.

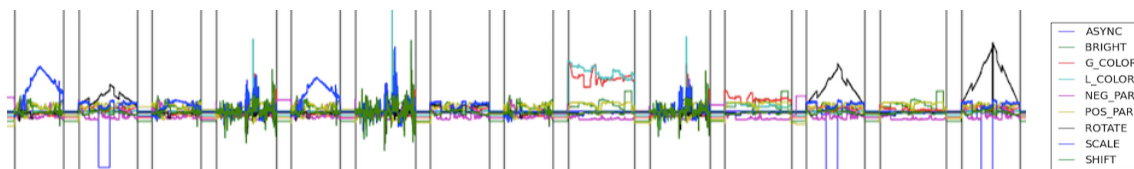


Рис. 15: Визуализация значений алгоритмов оценки качества для сцен тестового стереофильма. Каждая цветная линия соответствует измеренному определенному типу несоответствия между ракурсами, вертикальными линиями обозначены границы сцен.

Исходные данные качества стереовидео также подвергаются фильтрации: кадры с низким доверием не участвовали в обучении и тестировании. Доверие для кадра вычисляется с помощью информации о точности поблочного сопоставления ракурсов стереопары (алгоритм компенсации движения [10]). Таким образом, низким доверием обладают кадры с низкой яркостью, с

большим движением. Соответствующие таким кадрам значения объективной оценки усталости были исключены из выборок. Данные были нормированы следующим образом: 0 сопоставлялся со средним значением фрагмента выборки, а -1 и 1 соответствуют стандартному отклонению, взятому со знаком. Все каналы и мозговые ритмы ЭЭГ участвовали в обучении и тестировании независимо. В следующем разделе приводится точность работы разработанной системы.

4.4 Результаты

Разработанная система после обучения на данных проведенных экспериментов была протестирована на части полученных данных, не участвовавших в обучении.

На Рис. 16 визуализирована точность определения уровня дискомфорта зрителя от просмотра сцен, содержащих определенные артефакты. Каждая точка на графике соответствует одной сцене, по осям расположены измеренная в ходе эксперимента и предсказанная системой средняя величина усталости среди зрителей (Mean Opinion Score, MOS — усредненная среди данных всех людей оценка). Для измерения точности предсказания вычислялась ве-

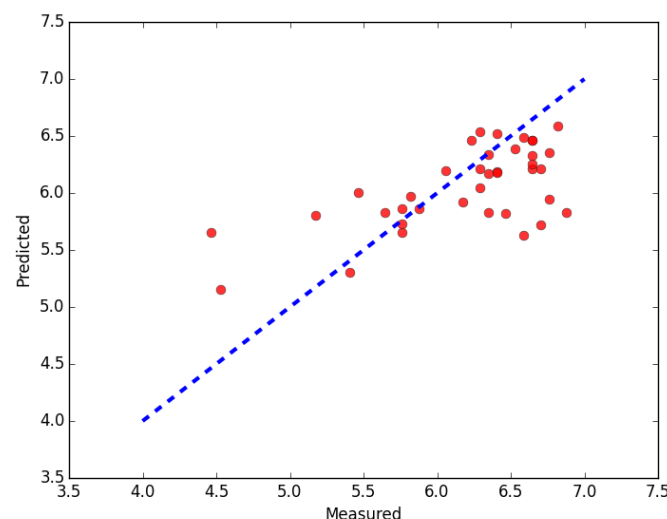


Рис. 16: Визуализация значений алгоритмов оценки качества для сцен тестового стереофильма. Каждая точка соответствует уровню усталости, определенному при просмотре 1 сцены стереофильма.

личина среднеквадратичного отклонения предсказанных значений усталости от измеренных:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (13)$$

Среднеквадратичное отклонение предсказанного значения от измеренного составило 0.612. При анализе относительного вклада признаков в построение решения обученной моделью на валидационных выборках выявлена закономерность, которая заключается в укрепившемся положении искажения временного несоответствия между ракурсами стереовидео на лидирующей позиции графика относительного вклада признаков в модель, показанного на Рис. 17.

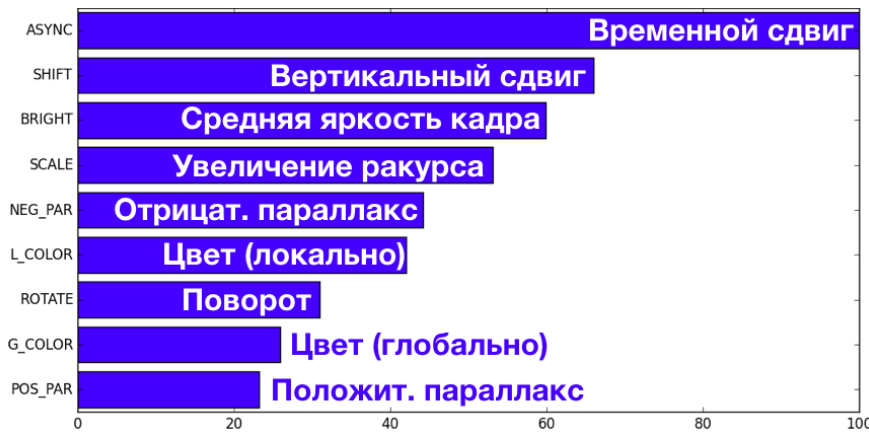


Рис. 17: Относительный вклад признаков, соответствующих различным несоответствиям ракурсов стереовидео в модель предсказания усталости зрителя. Самым значимым признаком при определении уровня дискомфорта выявлен временной сдвиг между двумя ракурсами стереовидео.

Как было упомянуто ранее, все каналы и ритмы анализировались независимо, и в ходе тестирования было выявлено, что наилучшим образом степень дискомфорта оценивается построенной и описанной в предыдущем разделе характеристикой альфа- и бета-ритмов. На Рис. 18 приводится точность работы построенной рекуррентной нейронной сети для альфа-ритма, на Рис. 19 — бета-ритма. На графике каждая точка соответствует значению дискомфорта, вызванного просмотром одного кадра стереовидео. При безошибочном предсказании график такого вида — это прямая $y=x$, где по оси X обозначено

измеренное (реальное) значение, а по оси Y — предсказанное. Так как количество кадров при обучении и тестировании равно нескольким десяткам тысяч, для удобства оценки распределения точек график построен в виде тепловой карты.

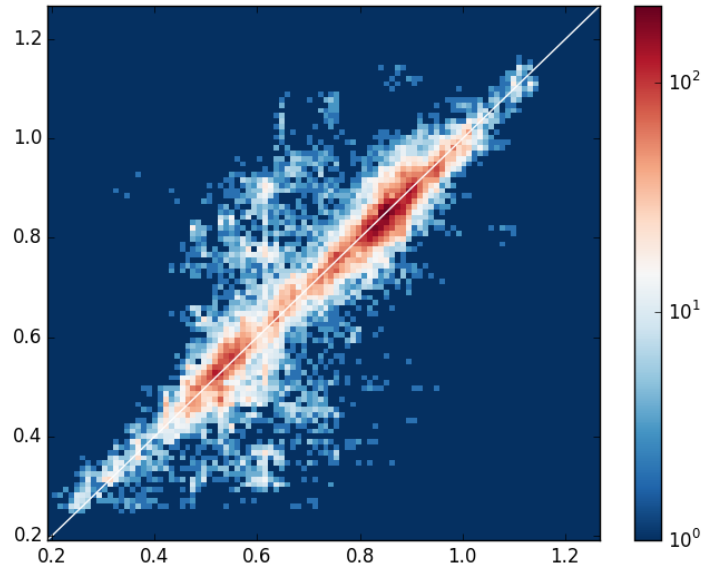


Рис. 18: Точность определения уровня дискомфорта, выраженного значением альфа-ритма, по данным качества стереовидео. График построен по данным канала №13 — Р4.

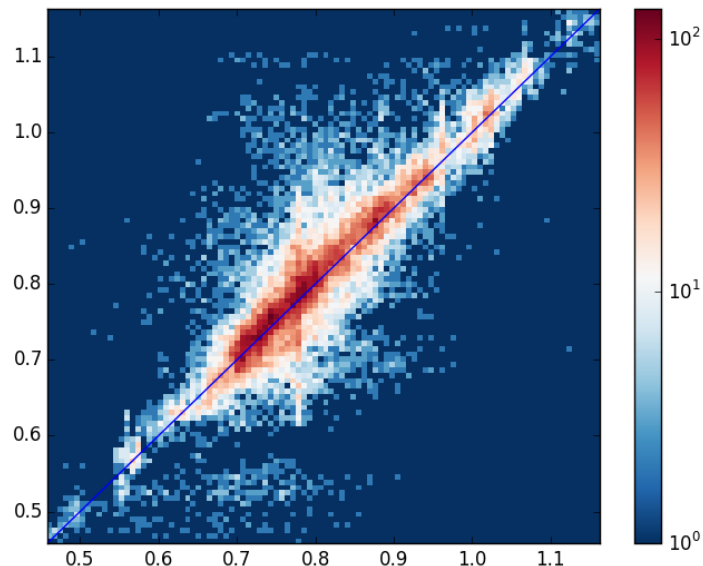


Рис. 19: Точность определения уровня дискомфорта, выраженного значением бета-ритма, по данным качества стереовидео. График построен по данным канала №13 — Р4.

На Рис. 20 и Рис. 21 изображены графики точности работы предложенного метода для всех каналов ЭЭГ, записанных во время объективного эксперимента.

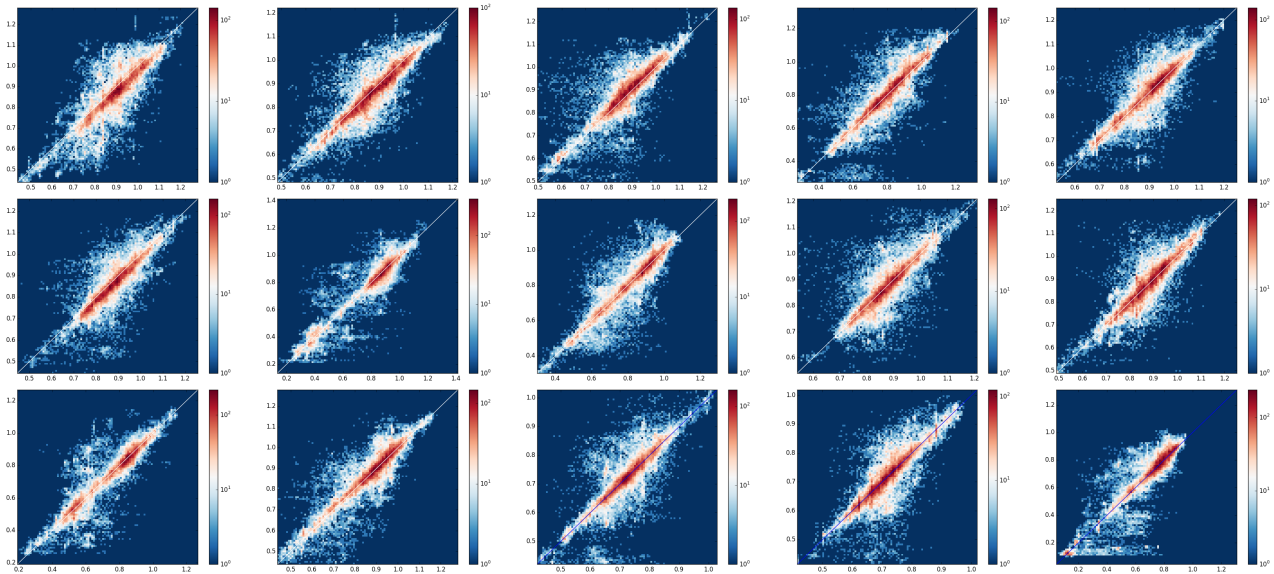


Рис. 20: Точность определения усталости на данных альфа-ритма.

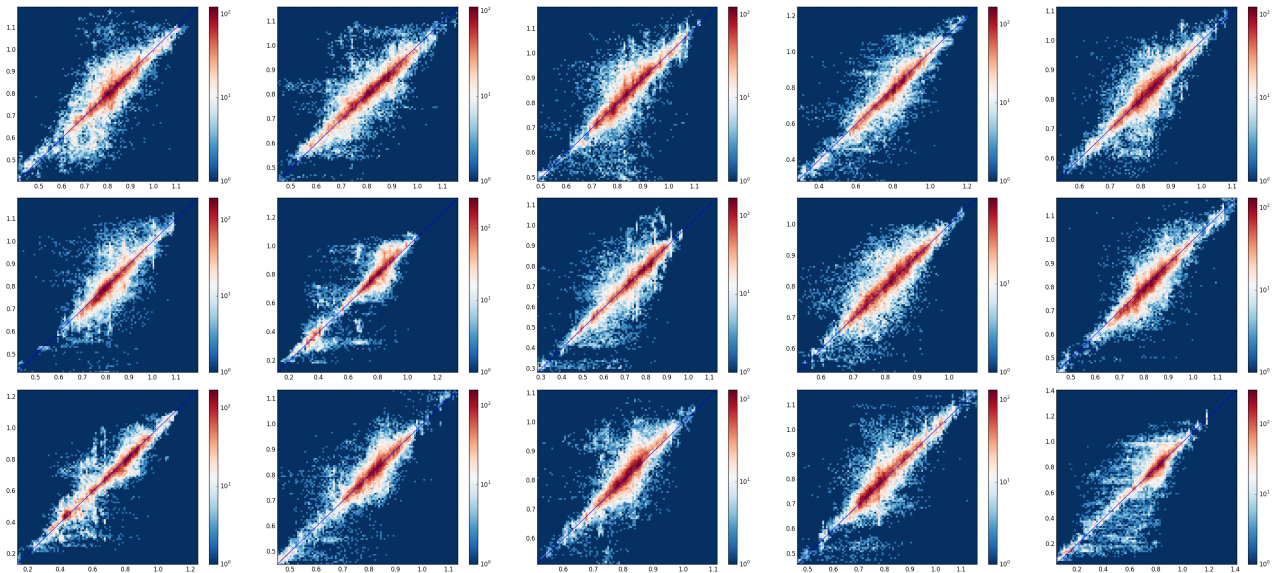


Рис. 21: Точность определения усталости на данных бета-ритма.

5 Описание практической части

Схема реализации метода для полученных субъективных данных об усталости зрителей представлена на Рис. 22. Схема реализации метода для полученных объективных данных (данных ЭЭГ) об усталости зрителей представлена на Рис. 23. Предложенный алгоритм для оценки уровня усталости от просмотра определенного стереофильма работает по схеме, изображенной на Рис. 24.

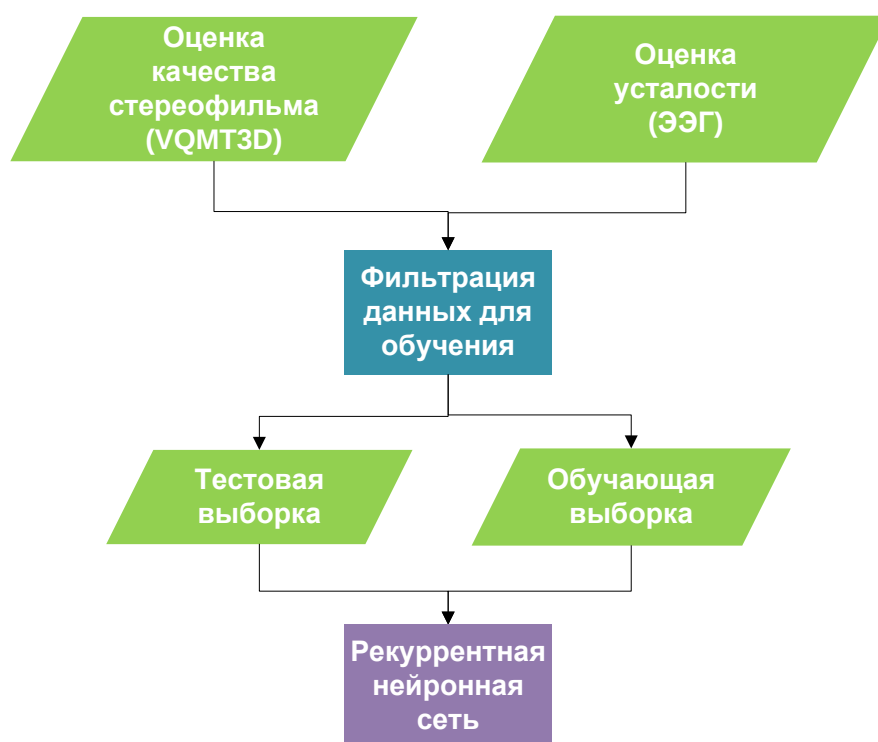


Рис. 22: Схема реализации алгоритма с использованием данных анкетирования.

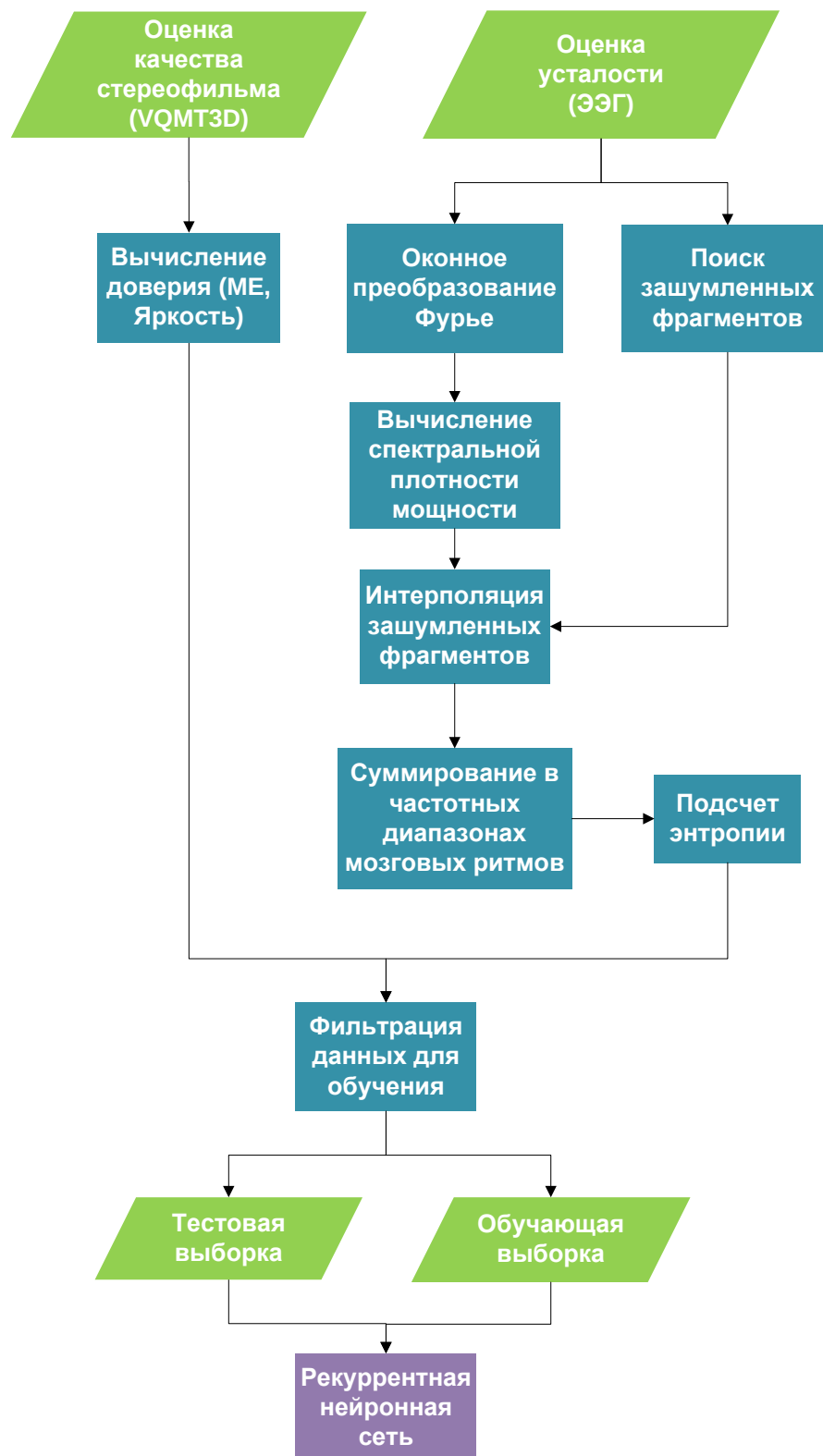


Рис. 23: Схема реализации алгоритма с использованием данных электроэнцефалограммы.

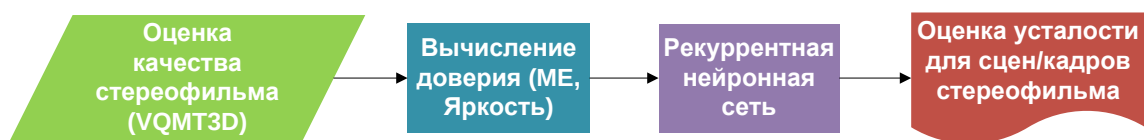


Рис. 24: Схема работы предложенного метода.

5.1 Генерация стереовидео для экспериментов

Для реализации описанного метода было проведено поэтапное разбиение на задачи. В первую очередь, для подготовки тестовых видеопоследовательностей в стереоформате необходимо было добавить искажения в исходные стереопары. Для преобразования ракурсов стереовидео использовался скриптовый язык AviSynth [2], средствами которого можно осуществить необходимые преобразования: поворот исходного видео; изменение масштаба; изменение частоты кадров, с помощью которого моделируется сдвиг во времени; добавление цветового искажения (с помощью модуля, разработанного видео-группой лаборатории «Компьютерной графики и мультимедиа»). Так как итоговое стереовидео содержит большое количество сцен с разными искажениями, программа на языке AviSynth состоит из множества схожих между собой по структуре и смыслу частей. Используя эту особенность, для генерации этого файла была разработана программа на языке Python, на вход которой подается закодированная в формате (1) последовательность, по которой она выдает файл в формате .avs для создания стереовидео. Программа на языке Python состоит из 250 строк; программа на языке AviSynth состоит из 1450 строк.

5.2 Обработка данных ЭЭГ

Для частотной фильтрации электроэнцефалограммы, а так же для вычисления показателей ритмов и их энтропии была разработана программа на языке C, состоящая в сумме из 800 строк.

5.3 Анализ качества стереовидео с помощью программ проекта Video Quality Measurement Tool 3D

Для оценки качества стереовидео использовались программные продукты, разработанные видеогруппой лаборатории «Компьютерной графики и мультимедиа» в рамках проекта VQMT3D [18]. Стереовидео, которые использовались для проведения эксперимента, были проанализированы алгоритмами оценки:

- величины положительной диспаратности - максимальное смещение объектов заднего плана дальше от зрителя относительно плоскости экрана
- величины отрицательной диспаратности - минимальное смещение объектов переднего плана ближе к зрителю относительно плоскости экрана
- величины цветового несоответствия ракурсов стереопары (локального и глобального)
- величины поворота одного из ракурсов стереопары относительно другого
- величины несоответствия масштаба ракурсов стереопары
- величины несоответствия между ракурсами стереопары по времени
- величины смещения по вертикали одного из ракурсов стереопары относительно другого
- средней яркости кадра

5.4 Обучение и тестирование искусственной нейронной сети

На имеющихся данных экспериментов была обучена искусственная нейронная сеть для решения поставленной задачи регрессии: определить число - степень усталости - по качеству стереовидео. Для реализации и построения этой сети был выбран язык Python и библиотека Keras [5]. Нейронная сеть имеет

3 рекуррентных скрытых слоя архитектуры LSTM: 300 нейронов в первом слое, 100 во втором и 10 в третьем. Функция активации - гиперболический тангенс. При обучении проводилось 100 итераций. Программа занимает 1500 строк. Тестирование обученной нейронной сети производилось на части данных эксперимента, не участвовавших в обучении. Результаты тестирования приведены в таблице Таблица 1.

Каналы	Ритмы				
	Альфа	Бета	Гамма	Дельта	Тета
Fp1	0.371	0.288	0.362	0.318	0.372
F7	0.355	0.308	0.486	0.413	0.301
F8	0.333	0.321	0.409	0.359	0.433
T4	0.354	0.209	0.240	0.288	0.303
T3	0.338	0.252	0.232	0.264	0.325
Fp2	0.333	0.303	0.304	0.314	0.301
O1	0.221	0.161	0.143	0.147	0.168
P3	0.290	0.226	0.214	0.247	0.234
F3	0.309	0.331	0.643	0.476	0.481
F4	0.306	0.201	0.171	0.211	0.198
C4	0.240	0.150	0.118	0.135	0.139
P4	0.268	0.211	0.329	0.235	0.191
POz	0.264	0.365	0.510	0.325	0.323
C3	0.234	0.257	0.320	0.271	0.256
O2	0.210	0.223	0.252	0.208	0.203

Таблица 1: Точность работы алгоритма: среднеквадратичное отклонение предсказанного значения от измеренного для исследуемых ритмов и каналов.

6 Заключение

6.1 Результаты работы

В данной работе предложен метод автоматического определения степени усталости зрителя, вызываемой просмотром стереоскопического видео. Разработанная система дает оценку дискомфорта, который испытает человек при просмотре стереовидео, содержащего различные несоответствия между ракурсами. Система использует объективные данные качества стереоскопического видео, содержащие информацию о наличии несоответствия между его ракурсами для каждого кадра.

Проведенное в ходе работы исследование электроэнцефалограммы головного мозга человека, записанной во время просмотра им подготовленного стереоскопического фильма с различными несоответствиями, позволила улучшить точность работы предложенной системы. Предложенный метод является первым в данной области методом оценки усталости зрителя, который работает с разрешением до кадра. В ходе исследования также было выявлено наиболее болезненное для зрителя несоответствие между ракурсами стереовидео - временной сдвиг между ракурсами. Несмотря на неестественную природу такого искажения, оно встречается в художественных стереофильмах.

6.2 Публикации и гранты

- Анциферова А. В. «Метод определения степени усталости, вызываемой просмотром стереовидео» // XXIII Международная конференция студентов, аспирантов и молодых ученых «Ломоносов-2016», стр. 8-10.
- Грант «Умник» (получен осенью 2015г.)

Также работа была частично поддержана грантом РФФИ №15-01-08632-а.

Список литературы

- [1] 3d cinema and headache: The first evidential relation and analysis of involved factors / Braschinsky Mark, Raidvee Aire, Sabre Liis et al. // *Frontiers in Neurology*. — 2016. — Vol. 7. — Pp. 1–6.
- [2] *AviSynthFAQ I*. Avisynth, 5 pages [online],[retrieved on dec. 7, 2007] // Retrieved from the Internet URL: <http://avisynth.org/medawiki/index.php>.
- [3] *Chen Wei*. Multidimensional characterization of quality of experience of stereoscopic 3d tv: Ph.D. thesis / Université de Nantes Angers Le Mans. — 2012.
- [4] *Cho Sang-Hyun, Kang Hang-Bong*. An assessment of visual discomfort caused by motion-in-depth in stereoscopic 3d video. // BMVC. — 2012. — Pp. 1–10.
- [5] *Chollet François*. Keras. — <https://github.com/fchollet/keras>. — 2015.
- [6] *Chuckravanen Dineshen*. Approximate entropy as a measure of cognitive fatigue: an eeg pilot study // *International Journal of Emerging Trends in Science and Technology*. — 2014. — Vol. 1, no. 07.
- [7] Classification and simulation of stereoscopic artifacts in mobile 3d tv content / Atanas Boev, Danilo Hollosi, Atanas Gotchev, Karen Egiazarian // IS&T/SPIE Electronic Imaging / International Society for Optics and Photonics. — 2009. — Pp. 72371F–72371F.
- [8] Eeg-based detection and evaluation of fatigue caused by watching 3d tv / Chunxiao Chen, Kun Li, Qiuyi Wu et al. // *Displays*. — 2013. — Vol. 34, no. 2. — Pp. 81–88.
- [9] Eeg-based mental fatigue measurement using multi-class support vector machines with confidence estimate / Kai-Quan Shen, Xiao-Ping Li, Chong-Jin Ong et al. // *Clinical Neurophysiology*. — 2008. — Vol. 119, no. 7. — Pp. 1524–1533.

- [10] Fast video super-resolution via classification / Karen Simonyan, Sergey Grishin, Dmitriy Vatolin, Dmitriy Popov // Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on / IEEE. — 2008. — Pp. 349–352.
- [11] A feature selection method for multilevel mental fatigue eeg classification / Kai-Quan Shen, Chong-Jin Ong, Xiao-Ping Li et al. // *Biomedical Engineering, IEEE Transactions on*. — 2007. — Vol. 54, no. 7. — Pp. 1231–1237.
- [12] *Kim Young-Joo, Lee Eui Chul*. Eeg based comparative measurement of visual fatigue caused by 2d and 3d displays // HCI International 2011–Posters' Extended Abstracts. — Springer, 2011. — Pp. 289–292.
- [13] *Liu Jianping, Zhang Chong, Zheng Chongxun*. Eeg-based estimation of mental fatigue by using kpca–hmm and complexity parameters // *Biomedical Signal Processing and Control*. — 2010. — Vol. 5, no. 2. — Pp. 124–130.
- [14] Movie Research Company. — <http://movieresearch.ru>.
- [15] Neurotrend. — <http://neurotrend.ru>.
- [16] *Series BT*. Principles for the comfortable viewing of stereoscopic three-dimensional television (3dtv) images. — 2014.
- [17] Ssvep and erp measurement of cognitive fatigue caused by stereoscopic 3d / Sungchul Mun, Min-Chul Park, Sangin Park, Mincheol Whang // *Neuroscience letters*. — 2012. — Vol. 525, no. 2. — Pp. 89–94.
- [18] Video Quality Measurement Tool 3D Project. — <http://compression.ru/video/vqmt3d/>.
- [19] Visual comfort assessment metric based on salient object motion information in stereoscopic video / Yong Ju Jung, Seong-il Lee, Hosik Sohn et al. // *Journal of Electronic Imaging*. — 2012. — Vol. 21, no. 1. — Pp. 011008–1.