

The image is a complex, surreal composition. At the top, a dark, starry space contains several planets and moons. A large, detailed Earth globe is the central focus, showing continents and clouds. This globe is reflected in a valley below, which features a river, mountains, and a forest. At the bottom of the image, a pair of large, blue, human-like eyes is superimposed over the landscape, looking directly at the viewer. The overall effect is one of a multi-layered, dreamlike world.

П. С. Варгин

**ОБЪЕМНОЕ
ТЕЛЕВИДЕНИЕ
И ЗРЕНИЕ**

ГОСУДАРСТВЕННАЯ КОРПОРАЦИЯ «РОСТЕХ»

Акционерное  общество

«Научно-исследовательский институт телевидения»



П. С. ВАРГИН

ОБЪЁМНОЕ ТЕЛЕВИДЕНИЕ И ЗРЕНИЕ

Санкт-Петербург
Издательство СПбГЭТУ «ЛЭТИ»
2025

УДК 621.397

ББК 32.94

B26

Варгин П. С.

B26 **Объёмное телевидение и зрение.** СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2025. 208 с.

ISBN 978-5-7629-3522-7

В книге впервые излагается концепция объёмного телевидения, обладающего свойствами глубины изображения и возможности для телезрителя действовать внутри изображения – интерактивности. Описаны принципы работы объёмной телекамеры и объёмного телевизора. Показано значение изучения свойств зрения человека для развития объёмного телевидения. Приведены результаты технического и математического моделирования различных особенностей зрения и слуха, позволяющие провести параллели между естественным и искусственным интеллектами, как средствами управления процессами отображения окружающей среды.

Книга содержит множество иллюстраций и предназначена для широкого круга инженеров, студентов и телезрителей, которых автор в книге именует теленавтами, подчёркивая их способность к интерактивности внутри виртуального мира – «мира телевизионного».

УДК 621.397

ББК 32.94

Рецензенты:

доктор технических наук, профессор **А. А. Гоголь**
(Санкт-Петербургский государственный университет им. М. А. Бонч-Бруевича);

кандидат технических наук, старший научный сотрудник **М. И. Курячий**
(Томский государственный университет систем управления и радиоэлектроники);

Институт электроники и телекоммуникаций Санкт-Петербургского политехнического университета Петра Великого (кандидат технических наук **В. А. Павлов**)

ISBN 978-5-7629-3522-7

© Варгин П. С., 2025

ПРЕДИСЛОВИЕ

Поскольку объёмное телевидение, как его понимает автор, пока не создано, назначение этой книги – помочь его созданию. Введение «объёмности» в многообразные применения телевидения улучшит качество видеоинформации, подобно введению цвета или увеличению чёткости изображения. Новое свойство телевизионного изображения должно и будет внедряться аналогично внедрению цветного телевидения – с соблюдением принципа совместимости имеющихся телевизионных приёмников с передаваемыми сигналами объёмного телевидения, хотя и с известными ограничениями.

Объёмное телевидение появляется вместе с предоставлением пользователю третьего измерения – дальности (глубины) изображения и возможности для телезрителя действовать внутри изображения (интерактивности). Природа, наши глаза и тело предоставляют нам такие возможности в реальном мире, а объёмное телевидение – в виртуальном. Создание объёмного изображения связано со стремлением инженера создать всё более полную иллюзию погружения телезрителя – по терминологии автора «теленавта» – в виртуальный (кажущийся реальным) телевизионный мир. Искусство создания иллюзии требует глубокого знания физиологии зрения, особенности которого должны быть познаны на уровне математических моделей. Такие модели могут создаваться копированием найденного свойства зрения, или путём выдвижения гипотезы о принципе работы зрительной системы с последующей её проверкой. Разработка гипотезы – это сложная синтетическая задача, требующая для своего решения не только знания физиологии зрения, но и знания технических методов обработки зрительной информации.

Автор книги в 1981 г. изобрёл телевизионный лазерный локатор (ТВЛЛ), т. е. активный способ определения пространственной формы объектов, который позволил получить компоненту дальности D телевизионного изображения. Вместе с R , G и B цветовыми компонентами появилась возможность создания активной объёмной $RGBD$ телекамеры. Сегодня компонента дальности называется картой глубины (*depth map*), а камера – времяпролётной (*Time of Flight, ToF*). Телевизионный лазерный локатор для телевидения был прямым выходом в третье измерение. В отличие от стереоскопической телекамеры новый способ давал информацию непосредственно о глубине изображения вне зависи-

мости от его оптического содержания, с максимальным разрешением по углам, дальности и времени.

Идеи автора спустя 30 лет после изобретения им телевизионного локатора были использованы в АО «НИИ телевидения» в ходе НИР «Ассистент» по разработке комбинированной – пассивной и активной – телевизионной системы орбитального обслуживания космических аппаратов.

Предшественником активного метода получения «карт глубины» с импульсной подсветкой инфракрасным излучением является широко известный пассивный метод анализа параллакса (диспаратности) элементов стереопары. Однако построение карт глубины пассивным методом затруднено поиском соответствующих точек изображений стереопары и необходимостью интерполяции значений карты глубины в промежуточных точках.

Карта глубины удобна для использования в системах технического зрения в сочетании с искусственным интеллектом и необходима при построении трёхмерной модели окружающей среды – мира телевизионного. Для наблюдения этого мира человеком следует представить на сетчатку его глаз проекции мира в виде изображений, которые соответствуют его текущей доминанте – направлению его взгляда. При этом объёмный монитор, например, наголовный дисплей виртуальной реальности, должен обеспечивать согласованную аккомодацию и конвергенцию глаз пользователя. К сожалению, современные стереодисплеи такой согласованности не обеспечивают, что обычно вызывает повышенную утомляемость зрения. Автор книги приводит описание не только ТВЛЛ, но и конструкции объёмного телевизора, обеспечивающего согласование аккомодации и конвергенции глаз телезрителя.

При построении глобальной (большой) сети объёмного телевидения, как и при строительстве сотовой телефонной связи, возникнут проблемы сжатия видеоинформации. Казалось бы, сжатие трёхмерной видеоинформации значительно сложнее обычной телевизионной. Но объёмное телевидение содержит ключи к сверхсильному сжатию, приближаясь к семантическому сжатию. Дело в том, что большинство объектов окружающей среды неподвижны, а подвижные объекты имеют малое число степеней свободы и могут имитироваться своими слегка подстраиваемыми аватарами (виртуальными копиями). Создание неподвижной сцены достигается разовым актом кодирования, например, путём сшивки изображений различных объёмных телекамер. Копия неподвижной,

сколь угодно большой, окружающей среды может храниться на достаточно ёмком сервере. Обращение к серверу возможно в части информации, выводимой на дисплей пользователя (телевизор «теленавта») в соответствии с его местоположением и ориентацией в телевизионном мире. Дополнительные возможности для сжатия дают особенности зрения человека. Направление взгляда может отслеживаться специальными трекерами и учитываться при предоставлении информации на объёмный дисплей наряду с ограниченной областью ясного зрения в пределах двух угловых градусов в направлении взгляда.

Сказанного про *объёмное телевидение*, по-видимому, достаточно для оправдания первой части названия книги. Что касается второй части названия, то следует заметить, что зрение человека содержит много загадок. Например, глаза все время совершают скачки и постоянно дрейфуют, но изображение внешнего мира для нас остаётся неподвижным (пространственная константность). Реализуется также цветовая константность вне зависимости от спектра источника освещения (лампы накаливания или дневного света). Отсутствие коротковолновых (синих) колбочек и палочек в зоне ясного зрения сетчатки глаза, отсутствие фоторецепторов в зоне слепого пятна не мешают нам воспринимать внешний мир без искажений. Наконец, передача в зрительную кору мозга лишь контурного изображения с сетчатки глаза, благодаря оппонентной организации рецептивных полей ганглиозных нейронов сетчатки, не мешает нам видеть интегральное изображение объектов наблюдения. Разгадка этих и других загадок зрения предлагается в этой книге.

Можно надеяться, эта книга поможет молодым читателям внести свой вклад в развитие объёмного телевидения с учётом физиологии зрения.

Доктор технических наук, профессор

А. К. Цыцулин

ОТ АВТОРА

Автор книги по образованию физик (радиофизик) в 1975 г. начал работу в ленинградском Всесоюзном научно-исследовательском институте телевидения (ВНИИТ). Исследования велись с большим размахом. Достаточно сказать, что в институте работало около 10000 сотрудников, из которых больше 150 в головном отделе научно-технической информации. Прекрасная техническая библиотека позволяла заказать любую книгу из фондов государственных библиотек. Большой штат переводчиков обеспечивал перевод любой иностранной статьи по заказу разработчика. Каждая пятница – день информации целиком посвящалась знакомству с мировыми новостями в области техники телевидения за неделю (эпоха Интернета еще не наступила). Число различных технических направлений телевидения в институте переваливало за десяток.

Интересны детали работы с молодежью: если молодой специалист сталкивался в дверях с руководителем лаборатории, то первым пропускался молодой товарищ, несмотря ни на какие возражения с его стороны. Директор института вставал со своего места и выходил на середину своего обширного кабинета, чтобы поздороваться за руку с впервые зашедшим к нему на приём молодым человеком. Этим в институте подчеркивалось – вся надежда в науке-технике на нас, молодых. Мы старались. В 1988 г. я защитил кандидатскую диссертацию под руководством д. т. н. профессора Л.И. Хромова.

Идиллия научной жизни быстро закончилась с наступлением «эпохи перемен» – перестройки в жизни России. С перекрытием финансирования исчезли НИР, резко сократился объём ОКР и на порядок снизилось количество сотрудников ВНИИТ. В этих условиях коллеги предложили мне создать акционерное общество, которое могло бы попытаться найти источник финансирования исследований по такому наукоёмкому направлению, как объёмное телевидение и таким образом избежать катастрофы – потери научного потенциала. В ноябре 1998 г. общество ЗАО «ПКБ «Рельеф» было создано. В его состав вошли четыре доктора и пять кандидатов наук. Как его генеральный директор, я стал писать статьи в журналы и письма с предложениями создания объёмного телевидения в различные инстанции. Инстанция на уровне Департамента цифрового телевидения Министерства связи вежливо

дала мне понять, что денег на науку нет и не предвидится. Акционерный капитал общества довольно быстро «съел» один из известных ныне банков, несмотря на отсутствие движения денежных средств по счёту предприятия.

Пришлось продолжить исследования в индивидуальном порядке. Добрых двадцать пять лет после изобретения ТВЛЛ (1981 г.) ушло на разработку концепции объёмного телевизора в обстановке «выживания», но патентовать и публиковать найденное техническое решение в тех условиях России было нецелесообразно. Во время исследований возникло понимание большого значения знаний о работе зрительной, а также слуховой (для создания объёмного звучания) систем человека. Причём

м знания эти должны быть сформулированы на языке математики для использования их в интересах инженерии. В книге приведены результаты работ по математическому моделированию некоторых частей этих систем.

За рубежом будущее объёмное телевидение развивается прежде всего в области виртуальной реальности. Одним из пионеров этой области является профессор С. М. ЛаВалль. Он написал прекрасную книгу с соответствующим названием (Steven M. LaValle VIRTUAL REALITY, 2020). Рекомендую русскоязычному читателю ознакомиться с моим переводом этой книги (см. на сайте telenavt.ru), выражаю надежду, что предлагаемая вниманию читателя книга автора наряду с книгой С.М. ЛаВалля поможет погружению в тему объёмного телевидения.

ВВЕДЕНИЕ

Заглядывая в будущее, следует помнить о прошлом. Мы помним, что телевидение развивалось от электромеханического (диск П. Нипкова) к электронному (кинескоп Б. Л. Розинга и иконоскоп В. К. Зворыкина), обретя на этом пути цвет (ЦТВ) и высокую чёткость (ТВЧ – 2К, 4К). Далее логично было использовать в стереотелевизионной системе (Д. Л. Бэрда) идею стереоскопа Эллиота–Уитстона–Брюстера–Холмса–Бейтса.



Пауль Юлиус Готлиб Нипков



Борис Львович Розинг



Владимир Козьмич Зворыкин



Джон Лоуги Бэрд

Стереоскоп эксплуатирует ряд свойств зрительной системы человека: анализ диспаратности, конвергенцию и аккомодацию глаз. К сожалению, обычный стереоскоп не обеспечивает надлежащего согласования аккомодации и конвергенции глаз, что приводит к их утомлению. Это обстоятельство и в настоящее время сдерживает широкое использование телевизионных стереодисплеев. В книге описан вариант пути решения этой проблемы для полной реализации свойства *объёмности* телевизионного изображения.

Отображающее устройство – дисплей – важная составная часть телевизионной системы, но система также должна включать датчик изображения. Для объёмного телевидения это объёмная телекамера, далёким прототипом которой является иконоскоп Зворыкина, а несколько более близким стереотелекамера Бэрда. Стереокамера даёт информацию о глубине поверхности наблюдаемого объекта в скрытом и неполном виде. Она не даёт непосредственно «карту глубины» объекта. С подобными тремя «картами» мы имеем дело в цветной телекамере. Это R , G и B (K , Z и C) компоненты цветного изображения. Карта глубины представляет собой матрицу сигналов глубины. Каждый элемент матрицы кодирует удалённость (глубину) соответствующего ему элемента объекта. В частности, сигнал глубины может квантоваться на 256 уровней, как и стандартные сигналы цветовых компонент. Карту глубины в реальном времени создаёт активная система с искусственной инфракрасной подсветкой – телевизионный лазерный локатор – ТВЛЛ (англ. *Time of Flight, ToF*) или пассивная с естественной подсветкой, реализующая карту диспаратности стереотелевизионной камеры, описанные в книге.

Объёмные $RGBD$ телевизионные камеры, где D – обозначение компоненты дальности (дистанции, глубины), содержат в удобном виде необходимую информацию для построения из их сигналов единого объёмного изображения, допускающего его оглядывание телезрителем. Возможность оглядывания объектов с разных сторон и движения между объектами виртуального телевизионного мира является проявлением свойства *трёхмерной интерактивности* – взаимодействия телезрителя с объёмным телевизионным изображением. Построение единого объёмного телевизионного мира производится путём шивки не только изображений объёмных телекамер различной угловой ориентации, но и изображений любых других датчиков локации окружающих объектов. Более того,

в объёмное изображение могут включаться и искусственные объекты (артефакты), реализуя тем самым дополненную виртуальную реальность.

Можно предположить, что большой интерактивный телевизионный мир в будущем будет строиться путём синтеза (сшивки) множества локальных виртуальных миров вплоть до выхода на глобальный уровень их объединения, подобно существующим сегодня *google*-картам. В сравнении с этими спутниковыми географическими картами построение глобального телевизионного мира потребует наличия более совершенных средств связи, поскольку телевидение предполагает передачу динамических сюжетов в реальном времени. Совершенствование средств связи тесно связано с эффективным сжатием видеоинформации. Казалось бы, сжатие трёхмерной видеоинформации значительно сложнее сжатия обычной телевизионной, но в действительности объёмное телевидение содержит ключи к особому сверхсильному сжатию. Дело в том, что большинство объектов окружающей среды неподвижны, а подвижные объекты имеют малое число степеней свободы и могут имитироваться своими слегка подстраиваемыми аватарами (виртуальными копиями). Создание неподвижной сцены достигается разовым актом кодирования. Хранящаяся копия неподвижной, сколь угодно большой окружающей среды хранится на достаточно ёмком сервере. Обращение к серверу осуществляется только для информации, выводимой на дисплей пользователя, в соответствии с его местоположением и ориентацией в телевизионном мире. Дополнительные возможности для сжатия дают особенности зрения человека. Область его ясного зрения (в пределах двух угловых градусов) в направлении взгляда отслеживается специальными трекерами и учитывается при предоставлении информации на его объёмный дисплей с учётом разрешения глаза по полю зрения.

Элементарная база будущего объёмного телевидения в настоящее время развивается по многим направлениям. Современные носители видеоинформации, например, КМОП фотоприёмники и видеосистемы на кристалле, с которыми мы сталкиваемся в своих смартфонах, продолжают интенсивно развиваться. Смартфон содержит не только компьютер, но и элементы навигационной системы, фото- и видеотехники, обеспечивает распознавание образов и перевод с одного языка на

другой, поиск информации в интернет с применением искусственного интеллекта. Самые современные образцы обладают даже *Lidar*-опцией, т. е. в них встроен ТВЛЛ. Пожалуй, наиболее близким к объёмному телевидению направлением развития видеотехники сегодня является направление создания систем виртуальной реальности.

Создавая трёхмерный телевизионный мир, человек идёт по пути копирования существующего физического мира, хотя и не ограничивается виртуальной копией, дополняя её по своему желанию. Но как начинающий скульптор, копируя работы предшественников, он должен пытаться постичь секреты их мастерства. Природа дала человеку зрение – своего рода объёмное телевидение, отображающее физический мир в его сознание. На создание зрения у природы ушли сотни миллионов лет в процессе эволюции Земли. Мы хотим создавать быстрее, пытаемся сжать время. Этого можно достичь, заимствуя опыт природы. Познать физиологию зрения – практическая необходимость для инженера. Познаёт инженер, используя свои инструменты и, прежде всего, математику. Пока нет математической модели явления, нет его понимания, нет возможности воспроизвести е

го в какой-либо форме. Физиологи, изучающие зрение, с большим трудом добывают информацию об отдельных его свойствах. Включить добытое в систему, чтобы прояснить свойства системы – осознать функцию нового в системе – очень трудное дело. Наше сознание для понимания ищет аналогии нового явления в своем опыте. Чем богаче опыт, тем больше ассоциативная база, и тем больше вероятность нахождения аналогии, то есть понимания. Искать аналогии непонятных процессов зрения в технике сложно ещё и потому, что сама техника быстро развивается, и те технические приёмы, которыми раньше владели многие «технари», сегодня ушли в прошлое, заслонились современными технологиями и теперь мало кому известны. Отсюда формула: новое – хорошо забытое старое. Полезно изучать историю техники и техники телевидения, в частности.

Непонятных процессов в зрении много. Десяток таких процессов перечислен в заключении к этой книге. Это те процессы, пониманию которых посвящена вторая половина книги.

1. Объёмное телевидение – символ XXI века?

XIX век – век пара и электричества,

XX век – век атомный,

XXI век – что станет его символом?

Многие изобретения, оказывая глубокое влияние на современную им науку и политику, изменяют мир и могут служить своеобразными символами эпохи. В XIX веке паровая машина привела в движение паровоз и пароход. Возможность за 80 дней (по расчётам Жюль Верна) обогнуть земной шар сильно уменьшила нашу планету в сознании человека. Телеграф и телефон способствовали ещё большему её уменьшению. Принцип относительности позволяет нам говорить о том, что произошло расширение нашего сознания. Атомный XX век сменил век пара и электричества. Возможность одномоментной гибели цивилизации от силы атомной бомбы обозначила ужасную перспективу – уничтожение сознания человека. Политикам удалось преодолеть возникшую опасность, подняв цивилизацию на новую ступень развития. Какое изобретение может стать символом XXI века? По-видимому, то, которое окажет наибольшее воздействие на наше сознание. Среди возможных претендентов на эту роль объёмное телевидение обладает преимуществом, поскольку оно по своей сути предназначено для этого. Есть основания предполагать, что воздействие может приобрести необычайную силу, следствием которой может стать захват сознания. Феномен захвата сознания состоит в невозможности или устойчивом нежелании сопротивляться постороннему воздействию на разум человека.

Каким будет объёмное телевидение? Может быть, Вы сразу представили себе стереотелевидение или стереокино, фильмы «Аватар» или «Алиса в стране чудес»? Это ошибочное, а точнее, далеко не полное представление.

С точки зрения телезрителя, по сравнению с обычным телевидением (плоским и стереоскопическим), объёмное будет обладать двумя новыми качествами: *трёхмерной интерактивностью* – возможностью взаимодействия с изображением, оглядывая объекты с разных сторон и двигаясь между ними и *объёмностью*, позволяющей глазам человека работать в естественном режиме, перемещая взор с близких объектов наблюдения на дальние. Ранее в литературе [1]–[4] обсуждалась возможность технической реализации объёмного телевидения и далее мы продолжим подробное обсуждение.

1.1. Анализ термина «объёмное телевидение»

Буквальный перевод слова «телевидение» – «дальновидение» – обозначает возможность наблюдения далёких от нас событий. Далёких, не только в пространстве, но и во времени, так как существует возможность записи событий прошлого. Заметим, что телевидение осуществляет также дальнослышание. В принципе, существует возможность передачи на большие расстояния запахов, тактильных и даже вкусовых ощущений. Возможно также дистанционное воздействие на вестибулярный аппарат человека. И всё-таки используется термин телевидение так как считается, что наибольший объём информации человек получает через зрение.

Слово «объёмное» созвучно слову «обнимать». Корень слова «ём» означает действие сходное со значением слов: «взятие, хватание». Это действие иллюстрирует рисунок 1, где объект А осуществляет «ём» объекта В. Сравним с фразой: Я «ем (кушаю)». Приставка «Об» или «о» означает круговой, всесторонний характер действия. Такое действие в заключительной его фазе иллюстрирует рисунок 2. На рисунках приведены плоские схемы, нам же следует представлять указанные действия как пространственные.

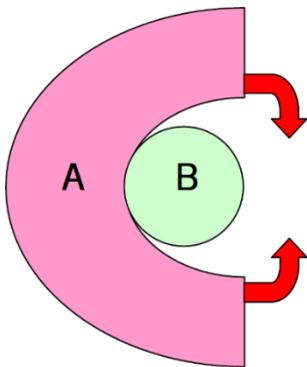


Рисунок 1 – «Ём»

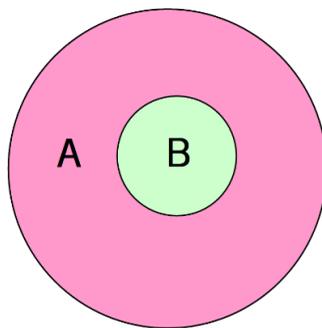


Рисунок 2 – Объём

Здесь мы говорим о телевидении и обнимать объекты собираемся в основном визуально, при помощи глаз и телекамер. Таким образом, в данном случае речь идёт о всестороннем охвате объектов физического мира. Использование терминов, связанных с телесными ощущениями, оправдывается их понятностью для всех носителей языка.

Иногда в качестве альтернативы термину «объёмное телевидение» предлагаются другие варианты. Например, стерео-

телевидение (от греч. στερεός – пространственный). Исторически термин привязан к методам создания, передачи и воспроизведения стереопары – двух слегка различных изображений для левого и правого глаза. Метод стереопары не подразумевает трёхмерной интерактивности и не обеспечивает полной объёмности изображения.

Другой вариант – телевидение со свободной точкой наблюдения (англ. – *free viewpoint television*) возник в связи с развитием многоракурсных методов воспроизведения плоских изображений. Этот термин обычно подразумевает ограниченную интерактивность (оглядывание), но не указывает на объёмность изображения.

Термин «виртуальная реальность» (от лат. *virtus* – потенциальный, возможный и лат. *realis* – действительный, существующий) – дословно «нереальная (кажущаяся, мнимая) реальность», в приложении к телевидению, акцентировал бы внимание на реалистичности воспроизведения модели мира телевизором. В телевидении качество телевизионного изображения, его реалистичность принято оценивать множеством предназначенных для этого характеристик. Говорить об уже состоявшейся реальности воспроизведения не приходится, поскольку нет предела совершенству. Вместе с тем, необходимо отметить, что термин «виртуальная реальность» в технике ассоциируется с многомодальностью (воздействием на различные органы чувств), что, конечно, соответствует направлению развития объёмного телевидения.

Переходя теперь к возможной классификации систем объёмного телевидения, хотелось бы заметить, что если Вы, уважаемый читатель, никогда не были в стереокино (3D-кинматографе), то вы можете все-таки представить себе, что такое объёмность телевизионного изображения, глядя в окно или в зеркало. Но если вам не довелось играть в компьютерные 3D игры, вам, скорее всего, будет трудно представить себе, что такое трёхмерная интерактивность. Важно не только посмотреть, как играют в «стрелялки» в ближайшем интернет-кафе, но и самому попробовать двигаться внутри изображения с помощью компьютерной мыши или джойстика. Трёхмерный телевизионный мир отличается от игрового мира тем, что последний нарисован художником – создателем 3D игры, а в случае объёмного телевидения он строится в реальном времени в памяти компьютера с помощью множества специальных объёмных телекамер. Опыт

трёхмерного взаимодействия с объёмным изображением на примере компьютерной игры позволит Вам с полным пониманием обратиться к последующему изложению.

Подчеркнем, что фундаментальное отличие системы объёмного телевидения от системы обычного плоского или стереоскопического телевидения состоит в том, что основным объектом телевизионной передачи является не сюжет, показанный телережиссером, а трёхмерный телевизионный мир, в котором разворачивается сюжет доступный наблюдению с любой, выбранной телезрителем, точки зрения, с задействованием всех механизмов восприятия этого мира зрительной системой человека. При этом и у самого телережиссера возможности формировать свою точку зрения на события резко возрастают. Пассивный телезритель может не включать режим интерактивности и смотреть на мир глазами режиссера. В любом случае при просмотре объёмной телепередачи появится захватывающая иллюзия реальности происходящего.

1.2. Классификация систем объёмного телевидения

Объекты наблюдения, которые составляют физический мир, доступны глазу человека и телекамере, благодаря формируемому ими световому полю. Обсуждая фундаментальные вопросы зрения человека, Э. Х. Адельсон и Д. Р. Берген в статье «Пленоптическая функция и элементы первичной стадии зрения» [5] ввели описывающую «всё, что можно увидеть» пленоптическую функцию (от лат. *plenus* – полный и англ. *optic* – оптический). Она представляет собой запись распределения интенсивности света внутри пучка лучей на входе оптической системы глаза $P = P(\theta, \varphi, \lambda, t, V_x, V_y, V_z)$, где: θ, φ – угол места и азимут линии визирования; V_x, V_y, V_z – координаты наблюдателя; λ – длина волны света; t – время. Эта функция позволила авторам создать таблицу базовых визуальных элементов, подобную периодической таблице Д. И. Менделеева. Элементы таблицы представляют собой двумерные сечения пленоптической функции. Некоторые из этих элементов являются математическими аналогами рецептивных полей зрительной системы человека. Функция нашла широкое применение в литературе по компьютерной графике и многокурсовому телевидению.

Для классификации систем объёмного телевидения можно использовать геометрическую часть пленоптической функции, при этом параметр λ , связанный с цветом излучения можно опустить, подразумевая, конечно, что объёмное изображение цветное. В отличие от глаза, объёмная телевизионная камера способна использовать зависимость пленоптической функции от времени для определения расстояния r от точки наблюдения до наблюдаемой точки объекта методом активной светолокации (см. ниже), при этом r и t связаны известной линейной зависимостью. Заменяем также оригинальные обозначения V_x, V_y, V_z на x, y, z , соответственно.

При новой параметризации получим: $P \equiv P(\theta, \varphi, r, x, y, z)$ – функцию шести переменных (измерений), где: θ, φ, r – составляют сферическую систему координат с центром в точке наблюдения; x, y, z – декартовы координаты точки наблюдения. Для определённости углы θ, φ можно отсчитывать относительно оси, параллельной оси глубины – z . Оси x – горизонтальную и y – вертикальную в одних случаях (для измерений) удобно выбрать параллельными плоскости окна объёмного телевизора, а в других (для навигации) привязать к какому-либо объекту телевизионного изображения. Шесть параметров (измерений) практически полностью определяют конфигурацию (геометрию) системы отображения в пространстве. Хотя в некоторых случаях полезным может оказаться введение параметра ψ – поворота телевизионной камеры вокруг оптической оси (крена), и тогда общее число измерений может вырасти до семи. Заметим, что у человека существует специальный механизм, парирующий повороты глаз вокруг оси зрения при наклоне головы вбок. Можно считать, что θ, φ, r – числовые значения координат точки поверхности наблюдаемого объекта в субъективной системе координат, x, y, z – объективные координаты центра зрачка глаза субъекта или объектива телекамеры.

Уточненный нами геометрический вид пленоптической функции позволяет говорить о шестимерности развитой системы объёмного телевидения. Учитывая это, можно составить классификацию систем объёмного телевидения, приведённую в таблице 1.

Таблица 1. Классификация систем объёмного телевидения

№ п/п	Параметры	Формула (полная)	Формула (сокращенная)	Телевидение
1	θ, φ	$2D(\theta, \varphi)$	$2D$	Двумерное (обычное)
2	θ, φ, r	$3D(\theta, \varphi, r); 2,5D(\theta, \varphi, r)$	$3D; [2,5D]$	Объёмное без оглядывания объектов наблюдения
3	θ, φ, r, x	$4D(\theta, \varphi, r, x)$	$4D(r, x)$	Объёмное с оглядыванием по горизонтали
4	θ, φ, r, x, y	$5D(\theta, \varphi, r, x, y)$	$5D(r, x, y)$	Объёмное с оглядыванием по горизонтали и вертикали
5	$\theta, \varphi, r, x, y, z$	$6D(\theta, \varphi, r, x, y, z)$	$6D(r, x, y, z)$	Объёмное с произвольным перемещением точки наблюдения
θ, φ – угол места и азимут линии визирования; r – расстояние от точки наблюдения до точки объекта; x, y, z – координаты наблюдателя				

Наш выбор параметров классификации систем объёмного телевидения соответствует геометрии отображения физического мира человеком и обусловлен стремлением к тому, чтобы телевизионная система максимально ей соответствовала.

Сегодня на рынке кинооборудования принято говорить о $3D$ (трёхмерном) стереоскопическом изображении. В $4D$ кинозале добавляются кресла с динамическими платформами. В $5D$ кинематографе добавляется комплекс спецэффектов (ветер, брызги воды, генераторы дыма и снега, имитаторы прикосновений и покалываний, инжекторы запахов). Такое употребление терминов ($3D, 4D, 5D$) обусловлено скорее рекламными, чем техническими соображениями.

В соответствии с приведённой таблицей можно выбрать формулу обсуждаемого варианта системы телевидения, например, $3D(\theta, \varphi, x)$, означает, что мы имеем дело с двумерным изображением θ, φ трёхмерной модели, при возможности её оглядывания, переме-

щаяся по горизонтали x . Такая формула отсутствует в таблице, содержащей лишь некоторые примеры. Смысл формулы очевиден. В данном случае в телевизионной системе отсутствует измерение r , то есть не предусмотрена возможность изменения аккомодации и конвергенции глаз при отслеживании объёмного сюжета.

Поскольку наличие координат θ и φ в телевизионной системе обязательно, их можно опускать и использовать сокращённую формулу, например, $5D(r, x, z)$. Последняя обозначает систему объёмного телевидения, или телевизора (формулы могут не совпадать) с трёхмерным (r) изображением и возможностью перемещаться в трёхмерном пространстве (трёхмерная интерактивность) в горизонтальной плоскости (x, z).

В нашей системе классификации обычная стереоскопическая система будет иметь обозначение $2,5D(r)$ или просто $2,5D$. Такое дробное обозначение нуждается в пояснении. Когда мы направляем свое внимание на определённую зону пространства, задавая дистанцию наблюдения, эта зона выделяется с помощью органа зрения. Выделение по углу происходит благодаря пику разрешающей способности глаза возле линии визирования, а выделение по глубине (дальности) – благодаря аккомодации и конвергенции глаз. Аккомодация (фокусировка) глаза обеспечивает ясное видение в пределах глубины резкости. Конвергенция (сведение) глаз минимизирует диспаратность на выбранной дистанции наблюдения, исключая двоение изображения. Если телевизионная система не обеспечивает функций конвергенции и аккомодации глаз, то в ней отсутствует измерение r ; если обеспечивает обе функции, мы говорим о наличии этого измерения; если обеспечивается только одна, то получаем 50 % реализацию измерения r или $0,5 r$. При наблюдении стереопары конвергенция глаз меняется при переносе внимания с близких точек объекта на дальние, а аккомодация остается неизменной, соответствуя дистанции расположения стереопары. Такая неестественная работа глаз приводит к их повышенному утомлению.

1.3. Схемы наблюдения

Вводя классификацию систем объёмного телевидения, необходимо отметить, что любая функция, включая пленоптическую, имеет область определения и область значений. В общем случае областью определения n -мерной функции с учётом возможного вырождения по координатам, является k -мерное

многообразие, где $k \leq n$, что подчеркивает возможную зависимость областей определения различных параметров пленоптической функции друг от друга. При оценке её области определения необходимо учитывать наличие двух миров: физического и телевизионного, системы пространственных координат которых необходимо привязывать друг к другу. Для случая классификации систем объёмного телевидения следует определять координаты точки наблюдения x, y, z в системе координат, привязанной к телевизионному миру (пространству телевизионного изображения).

Рисунок 3 иллюстрирует «аквариумную» – с замкнутым объёмом изображения – схему объёмного телевидения, когда наблюдаемый сюжет ограничен сферической поверхностью. В данном случае систему телевидения можно характеризовать формулой $6D(\theta, \varphi, r, x, y, z)$. Точка наблюдения здесь находится в точке физического пространства, внешней по отношению к пространству изображения.



Рисунок 3 – Объёмный телевизор замкнутого (аквариумного) типа, замыкает (вырезает) небольшой сферический объём большого телевизионного мира

В отличие от физического аквариума Вы можете менять с помощью джойстика содержание объёмного изображения в интерактивном режиме, виртуально изменяя координаты x, y, z и перемещаясь, таким образом, в совершенно другую область телевизионного мира. С другой стороны, Вы можете, при желании, обойти свой «аквариум» кругом, приблизиться или удалиться от него, тем самым, физически изменяя свои координаты x, y, z в физическом мире. В обоих случаях центр системы координат x, y, z остается привязанным к какой-либо выбранной точке объёмного изображения.

Приведённую схему наблюдения большого телевизионного мира трудно назвать естественной, поскольку мы привыкли видеть окружающий мир без резких ограничений его объёма. Конструкция монитора (телевизора) должна обеспечивать восприятие объёмного изображения без искусственных ограничений по глубине и с достаточно широким полем зрения. Рисунок 4 иллюстрирует схему объёмного телевидения с телевизором открытого типа.



Рисунок 4 – Объёмный телевизор открытого типа; обеспечивает естественный обзор телевизионного мира из малого виртуального объёма

В этом случае у телезрителя легко возникает иллюзия присутствия внутри телевизионного изображения, при минимально ощущаемых собственных размерах. Здесь мы также имеем дело с системой телевидения, характеризуемой формулой $6D(\theta, \varphi, r, x, y, z)$. Центр системы координат x, y, z в данном случае также следует привязать к выбранной точке объёмного изображения. Для обеих схем наблюдения может оказаться удобным поворот системы координат x, y, z вокруг вертикальной оси y так, чтобы ось z была направлена в сторону наблюдателя.

Область пространства изображения в совокупности с привязанной к ней областью физического объёма возможного перемещения точки наблюдения представляет собой пространство наблюдателя.

Область определения координат θ , φ , r , зависящая от положения точки наблюдения и поворота зрителя представляет собой пространство (зону) наблюдения.

Все познается в сравнении. Можно было бы вести рассказ об объёмном телевидении, сравнивая его с обычным «плоским» телевидением, но более экономным представляется путь проведения параллелей между системой объёмного телевидения и сознанием человека. Сознание – субъективная модель окружающего мира, формируется в процессе активности человека в трёхмерном физическом мире. Всю жизнь мы совершенствуем эту модель, познавая мир.

1.4. Когнитивная наука

Весь комплекс таких наук, как психология, лингвистика, нейрофизиология, философия, искусственный интеллект, нейроинформатика, антропология, экономика и др., предоставляющих свои методы и наработки для исследований познавательных возможностей человека принято сегодня называть когнитивной наукой. Можно предполагать, что её достижения будут востребованы объёмным телевидением в полной мере, оно же, в свою очередь, будет влиять на развитие этой науки.

При создании развитой системы объёмного телевидения человечество вплотную подойдет к тайне сознания, важнейшим инструментом которого является воображение. Наше воображение воспроизводит объёмные зрительные образы. В памяти человека содержится неисчислимое их количество. Слова нашего языка являются кодом доступа к этой памяти. В своем сознании мы можем масштабировать, вращать, перемещать объёмные сцены, синтезировать их цепочки. Мы можем определять жизнь действующих лиц, живущих в наших мыслях, многократно проигрывая и изменяя историю их бытия. При этом воображаемые объекты обычно похожи на реальные, но могут приобретать иногда и совершенно необычные формы. По существу, реальный мир – пища для нашей фантазии.

Объёмное телевидение будет обладать всеми возможностями воображения отдельного человека и превосходить его, поскольку оно, в части информационного содержания, будет плодом усилий всей цивилизации.

До изобретения телевидения и фотографии человек отправлялся в путешествие, смотрел по сторонам, возвращался

домой и рисовал соседям картины дальних стран с помощью карандаша или словами с помощью воображения собеседника. Фотоаппарат, а затем киноаппарат сделали эти картины более достоверными. Телевидение ликвидировало задержку во времени. Большинство людей знакомы с фото-, кино- и телетехникой. Многим приходилось играть в компьютерные 3D игры. При этом, далеко не каждый, по-видимому, задумывался над тем, как он осознает окружающий мир. Однако для того, чтобы заглянуть в будущее и представить себе возможности объёмного телевидения необходимо задуматься и над этим, приготовившись к погружению в синтез компьютерных, телевизионных и когнитивных технологий. Говоря о сознании человека, мы должны представлять себе, что это просто весьма изощренный механизм отображения реальности. Инженеру, занимающемуся разработкой системы объёмного телевидения, которая также представляет собой механизм отображения, следует освоить особую не техническую терминологию для использования в своем творчестве технических решений природы.

Поскольку сознание формируется в процессе познания, предельно упрощенная модель процесса познания будет полезна при обсуждении свойств объёмного телевидения. Вернёмся к началу изложения. На рисунке 1 тело A познаёт (осязает) поверхность тела B и его форму. Объект B «знал» (сохранял) свою поверхность и форму. Субъект A познал (скопировал при контакте) эти свойства B и таким образом возникло *со-знание* A , то есть у A появилось совместное с B знание. Так наглядно мы определили две важнейших функции сознания: познание как процесс и знание как результат этого процесса.

Субъект (носитель сознания) A кроме B поочередно познает множество объектов C, D, \dots окружающей его среды. Познав объекты по отдельности, субъект познает их как множество, выявляя их соотношения. Соотношения могут динамично изменяться.

Человек слепой и глухой познает предметы окружающего мира в основном через осязание или объятие. Кроме относительных координат элементов поверхности предмета он может определить её шершавость, твёрдость, теплоту, влажность и пр. Зрячий и слышащий человек активно пользуется дистанционным методом познания окружающей среды. Именно так можно быстрее «объять необъятное».

Субъект познает необъятный (превосходящий размер его объятий) мир, моделируя этот мир. Он создаёт модель окружающего мира в своем сознании из фрагментов своих ощущений. Информация зрения, слуха, осязания автоматически наносится на некую «карту», которая постоянно, при смене ракурса наблюдения, сверяется с окружающим миром и уточняется. При постоянных движениях глаз и тела человека для привязки карты к местности важен учёт этих движений. Так обеспечивается константность восприятия человеком пространства.

В рамках данной темы нас не слишком интересует, какова физическая реализация этой карты (памяти), то есть, представляет ли она собой нейронную или синаптическую сеть, или биохимический комплекс, также как пока нас не интересует, будут ли реализованы схемы объёмного телевидения с применением фотолиитографии или каких-либо новых нанотехнологий.

В математике существует понятие гомеоморфизма – взаимно однозначного соответствия между двумя математическими объектами. Если модель мира субъекта соответствует, как нам кажется, реальности, мы называем её адекватной, но при этом мы должны понимать, что модель никогда не бывает гомеоморфной. Доказательством последнего является наличие оптических иллюзий (обмана зрения). Например, на рисунке 5 многие люди, хотя и не все, видят вращение колес, которое на самом деле отсутствует.

Восприятие мира может сопровождаться иллюзиями, связанными с достройкой модели мира, при отсутствии необходимых данных от органов чувств, то есть с фантазией человека и его воображением. В языке это свойство сознания закреплено в выражении «мне показалось». Правильная достройка возможна потому, что человек знает законы изменения формы мира при перемещении в нём. Эти законы он постигает ещё в детстве во время игр. Учитывая эти законы, сознание субъекта может не только достраивать модель мира в процессе восприятия, но и создавать собственные воображаемые миры.

Телевизионная модель физического мира, создаваемая из фрагментов, получаемых с помощью объёмных телекамер, также должна обладать целостностью и стабильностью воспроизведения (константностью). Во многих случаях мы можем не добиваться идентичности телевизионного мира миру физическому, ограничившись их похожестью.

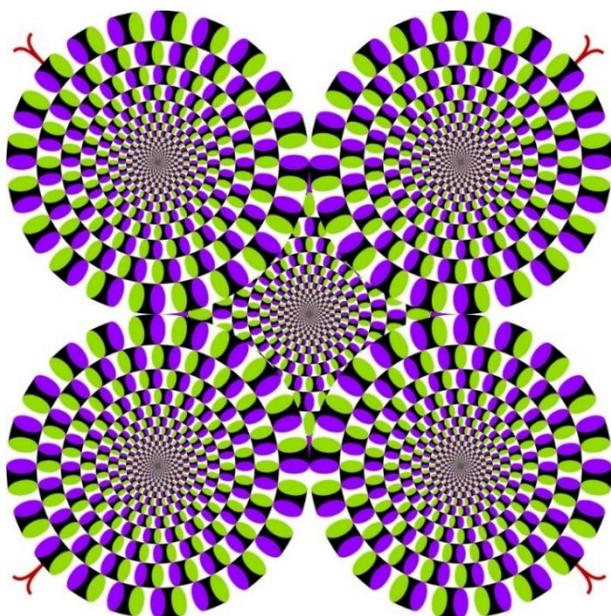


Рисунок 5 – Иллюзия вращающихся колёс

Объёмное телевидение в значительной мере будет копировать как окружающий мир, так и сознание человека. Творческий коллектив, участвующий в создании объёмной телепередачи, может создать телевизионную модель мира, в которой могут действовать персонажи реального мира, анимационные персонажи, а также и телезрители. Такую модель можно назвать смешанной реальностью (англ. *mixed reality*). Другими словами, телевизионный мир в известном смысле сложнее физического, поскольку может представлять собой синтез моделей физического и воображаемых миров.

1.5. Три мира

Так устроен человек и так идёт развитие цивилизации, что в будущем мы сможем иметь доступ в три трёхмерных мира: физический, телевизионный и мир воображения. Каждый из этих миров допускает *трёхмерную интерактивность*.

На рисунке 6 символически показаны эти три мира: физический – 1, телевизионный – 2, и воображаемый – 3, – воспринимаемые (сплошные стрелки) субъектом 4 в соответствии с переключением его внимания (пунктирные стрелки). Миры связаны между собой взаимодействием – интерактивностью (двунаправленные стрелки).

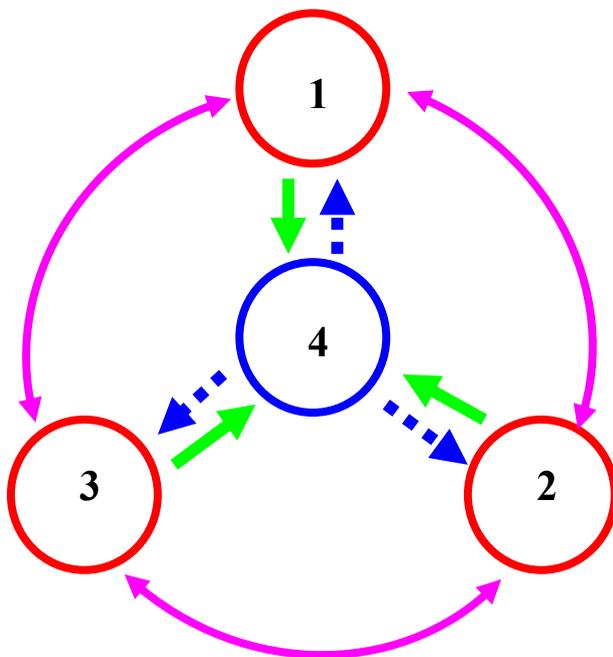


Рисунок 6 – Наблюдаемые миры: 1 – физический, 2 – телевизионный, 3 – воображаемый; субъект 4 воспринимает то (сплошная стрелка), на что направлено его внимание – стрелка (пунктирная)

Поскольку для субъекта важно не просто увидеть, а узнать объект наблюдения, то его восприятие происходит одновременно с действием его воображения, то есть одновременно с синтезом умственной модели объекта путём трансформации содержимого памяти субъекта. Если субъект наблюдает мир 1, то его внимание переключается с высокой скоростью, от впечатления из мира 1 к впечатлению из мира 3. Субъект подобен скульптору, многократно переводящему взгляд с живой модели на мраморную, которую он должен сделать подобной оригиналу. Аналогично, при наблюдении телевизионного мира внимание переключается между миром 2 и 3. Аналогия зрительной модели со скульптурой весьма условна, поскольку зрительная модель – это лишь некоторый код, извлекаемый из памяти субъекта, который создаёт у него иллюзию видимости знакомого объекта. Эта иллюзия–гипотеза позволяет ему направлять взор на информативные (характерные) участки модели и сличать их с участками видимого внешнего мира, реагируя

на различия. Если гипотеза подтверждается, он решает, что узнал объект внешнего мира, если нет – выдвигает новую гипотезу, если объект ни на что не похож, конструирует из знакомых элементов синтетический образ, называет его и запоминает для будущих узнаваний. Поскольку время пространственной фиксации взгляда ограничено периодом между произвольными скачками глаз – микросаккадами (около трёхсот миллисекунд), количество последовательных приближений модели к впечатлению от объекта наблюдения ограничено примерно десятью итерациями. Обычно человек быстро и хорошо узнает объекты внешнего мира, что говорит о весьма эффективном кодировании и обработке визуальной информации в его зрительной системе. Мы можем узнавать объекты по нескольким штрихам их рисунка, размытой чёрно-белой фотографии, единственному, характерному цветному пятну или особенности движения. Это наводит на мысль о различных, взаимно дополняющих системах кодирования. Действительно, в коре головного мозга найдены [6] специализированные нейроны, реагирующие на направления штрихов, углы, пространственные частоты, цвет и движение наблюдаемых объектов.

В развитой системе объёмного телевидения при отображении мира 1 в мир 2 процесс отображения может происходить во многом аналогично. Распознанные объекты могут сличаться с их моделями, что выгодно с информационной точки зрения. При этом с каждым телевизионным кадром имеющаяся большая телевизионная трёхмерная модель мира не должна обновляться целиком, а лишь корректируется в той её части, где за время кадра появились различия между моделью и отображением внешнего мира. Телевизионная система, состоящая из множества объёмных телекамер, может следить за физическим миром и обращать своё внимание только на существенные его изменения. Реагирование на изменения – общий принцип живых организмов, выработанный долгой эволюцией.

При обучении синтезу умственной модели ребенок совершает её аффинные (масштаб, поворот, смещение) преобразования, например, при игре с кубиком, а также преобразования сжатия–растяжения–изгиба при игре с мягкой или резиновой игрушкой. Заметим, что ввиду особенностей зрения младенца, поначалу все доступные ему трансформации происходят в области низких пространственных частот и лишь с возрастом он осваивает высокие

пространственные частоты, расширяя пространственную зону наблюдения, а к полугоду развивает бинокулярное (трёхмерное) зрение.

В системе объёмного телевидения могут осуществляться практически любые трансформации телевизионного мира как на передающей, так и на приёмной стороне. Можно, например, осуществить инверсную стереографическую проекцию телевизионного мира в сферу по формуле $\rho = R^2/\rho'$, где: ρ и ρ' – радиальные координаты точек внешнего и внутреннего пространства сферы с началом в её центре, а R – радиус сферы, равный, например, радиусу сферы объёмного телевизора (рисунок 4). В этом случае гигантский внешний мир «свернётся» в объём сферы. Разумеется, при этом он претерпит громадные нелинейные искажения, но он будет вполне точной объёмной картой этого мира, а мы сможем в интерактивном режиме проникнуть внутрь этой карты.

При освоении пространства физического мира с первых дней жизни человек кроме зрения активно, при наличии хватательного и сосательного рефлексов, используется осязание. Ребенок осваивает мир, обнимая его в прямом и переносном смысле. Постепенно он формирует свое перцептивное (ощущаемое) пространство, которое и является его моделью наблюдаемого физического мира. Заметим, что геометрические свойства этого пространства отличаются от свойств евклидова пространства. Эти свойства изучаются как экспериментально, так и путём анализа геометрических свойств картин, нарисованных различными художниками, в которых явно присутствуют законы перцептивной проекции. Математическое описание геометрических свойства перцептивного пространства формируемого из сетчаточного образа путём растяжений и сжатий его элементов под действием механизмов константности величины и формы получено, например, в работе [7].

Здесь уместно задать себе чисто телевизионные вопросы. Каково качество изображения субъективного мира при закрытых глазах, то есть при отсутствии синхронизирующих воображение внешних впечатлений? Насколько чётко может воспроизводить мир наше воображение? Насколько насыщенными могут быть цвета воображаемых трёхмерных объектов? Какова динамика этих объектов? И, наконец, каковы возможности трёхмерной интерактивности в воображаемом мире? Вопросы эти отнюдь не праздные, поскольку обсуждаемое здесь взаимодействие трёх миров может приводить к самым неожиданным последствиям для человека.

Если мы закроем глаза и постараемся вообразить цветную трёхмерную сцену, вряд ли эта попытка нас удовлетворит. Скорее всего, мы зафиксируем что-то вроде фрагмента из мира теней. Яркое – эйдетическое воображение, когда человек в течение нескольких минут способен удерживать перед закрытыми глазами образ увиденного предмета, по наблюдениям психологов, редкость и чаще встречается у детей в возрасте до 5 лет.

Особым состоянием сознания является сон. Во сне мы сталкиваемся с действием нашего воображения при отсутствии внешних впечатлений из-за «отключенных» органов чувств. Все люди видят сны, но не все их помнят. Помнящие сновидцы иногда описывают очень яркие и чёткие цветные сны. Исключительный интерес представляет динамика сна. Сновидения, по наблюдениям учёных, возникают в фазе быстрых движений глаз. В этой фазе сна, появляющейся несколько раз за ночь, человек, видящий сон, переводит взор с одних объектов мира сна на другие, при этом его глаза физически совершают скачки (саккады), аналогично тому, как это происходит наяву. То есть глаза сохраняют подвижность, в то время как все остальные части тела спящего человека бездвигательны. Это свойство глаз в принципе позволяет осуществлять объективный контроль визуальных впечатлений спящего человека, например, методами электроокулографии. При этом, по-видимому, можно построить картину движения глаз спящего, подобную приведённой в работе А. Л. Ярбуса [8] для бодрствующего человека, показанную на рисунке 7, когда след движения глаз позволяет угадать объект наблюдения.

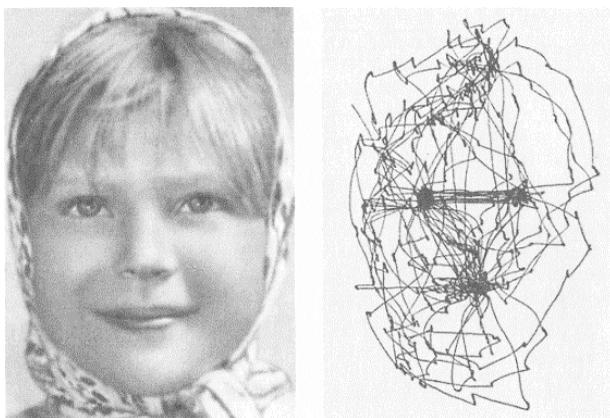


Рисунок 7 – Слева рассматриваемый портрет.
Справа запись движений глаз

Как правило, внимание сновидца не принадлежит ему самому. Он видит сон как пассивный кинозритель, по программе не им созданной. Являясь персонажем сна, он следует за его сюжетом некритически, оправдывая для себя самые нелепые его повороты. Такие сны вспоминаются обычно (но не обязательно) как почти чёрно-белые, не слишком чёткие иллюстрации к различным эмоциональным состояниям. Это прекрасный пример захвата сознания, в данном случае «самозахвата».

Совершенно иная картина наблюдается в осознанном (люцидном) сновидении. Спонтанно осознанные сновидения возникают редко и, по-видимому, не у всех. Человек видит сон и знает, что спит. Во сне он обладает волей, способен формулировать и выполнять свои желания, ставить эксперименты и при этом помнит о решениях, принятых им наяву. Методики обучения переходу к осознанному сновидению, разработанные, например, в Центре изучения сна при Стэнфордском университете, позволяют обучить ему любого человека. Особенно важно то, что движения глаз *онейронавта*, фиксируемые с помощью полиграфа («детектора лжи» в некоторых приложениях) позволяют сновидцу установить одностороннюю связь с лабораторией. Начиная и заканчивая эксперимент во сне, он подает сигнал об этом заранее оговоренным движением глаз, например, «влево – вправо» несколько раз.

Осознанно сновидящий прекрасно помнит свой сон после пробуждения, его сон цветной, яркий, четкий, не уступающий, при желании, по качеству изображения наблюдению физического мира наяву. Он способен силой своего желания изменять динамику сновидения, то есть обладает практически любой трёхмерной интерактивностью, включая полёты на заданной высоте и осуществление «материализации» предметов. Судя по многочисленной литературе мистической направленности, в мир осознанного сновидения можно перейти и после соответствующего обучения, минуя фазу засыпания, путём погружения в глубокую медитацию.

1.6. Захват сознания

Человек общественное существо. Общество, составной единицей которого он является, может потребовать от него действий, не совместимых с его личными интересами. Это требование может быть продиктовано интересами выживания социума, в который входит данный человек. Например, государство

может отправить его воевать и заставить постоянно рисковать своей жизнью. По-видимому, на протяжении бесчисленных веков эволюции у человека сформировался физиологический механизм, помогающий его сознанию перейти в ведомый режим и действовать без лишних страхов и сомнений, выполняя приказы. Похоже, что такой режим сознания мы и наблюдаем в обычном сновидении.

С другой стороны возможности выживания отдельной личности и общества, в котором она привычно существует, зависят от развития и самостоятельности самой личности. Герой – личность наивысшей степени общественной ценности – не только смел, но и умел. Заметим, что в осознанном сновидении человек, сознавая что он спит, обладает смелостью, собственной волей и умением решать возникающие задачи с помощью своей мысли.

Стремительное, по меркам исторического развития, вхождение в мир человека телевидения может нарушить сложившийся за его долгую эволюцию баланс между активностью и пассивностью, умелостью и беспомощностью, индивидуализмом и альтруизмом. Конечно, у человечества есть опыт изменения баланса. Во время войны доминируют социальные отношения, а в мирное время выходит на первый план частная инициатива, но при этом почти не затрагиваются репродуктивные основы цивилизации. Люди всегда учили своих детей тем навыкам, которые сами обрели в детстве. Во всяком случае они, как правило, не мешали детям развиваться.

В предыдущем разделе мы касались вопроса формирования перцептивного пространства ребенка. Модель мира надо научиться строить в своем сознании, исследуя мир комплексно с помощью всех органов чувств, а вовсе не только с помощью зрения. Если подвергнуть живое существо депривации (лишению какого-либо информационного канала) в процессе его интенсивного развития (в детстве), то в дальнейшем никакими усилиями невозможно восстановить нормальное функционирование подвергнувшегося блокировке органа чувств. Если посадить ребенка перед телевизором, вместо того чтобы ежедневно читать ему вслух, объясняя прочитанное, играть с ним в игрушки, петь, прививать интерес к поделкам, то последствия будут весьма плачевными. Исследования физиологической активности мозга «теледетей», проводящих перед телевизором ежедневно по десять–пятнадцать часов, показывают, что у них «кора головного мозга подобна пустыне». «Они страдают

полной *потерей способности воображения*». «Некоторые дети даже не в состоянии нарисовать по памяти бытовые предметы, скажем чашку» [9].

Потеря произвольного воображения – вот что угрожает телеману! То есть, согласно рисунку 6, мир 2 (телевизионный) может вытеснить мир 3 (воображение). Человек может также потерять вкус к активной жизни. То есть мир 2 может полностью заменить мир 1 (физический) в нашем сознании. Возникает, если можно так выразиться, «война миров». Общество, с потерей навыков и интереса к активной жизни своих членов, слабеет, а человек, с уничтожением самостоятельного воображения теряет способности творца. А между тем, формула счастья звучит так: «Сам придумал и сам воплотил».

Когда мы говорили о телемании, то есть о навязчивом стремлении смотреть телевизор, то имели в виду уже существующее двумерное телевидение. О воздействии телевидения на детей можно прочесть в книге немецкого ученого и учителя Р. Пацлафа «Застывший взгляд» [9].

Объёмное телевидение, как уже говорилось во введении, обладает двумя новыми качествами: *объёмностью* и *трёхмерной интерактивностью*. Эти качества усиливают иллюзию реальности происходящего в окне объёмного телевизора, понижают порог критического отношения телезрителя к происходящему и сознание человека с большей легкостью может переходить в ведомый режим, когда происходит его захват.

Надо сказать, что свойство объёмности, само по себе весьма полезно для глаз человека, поскольку он не вынужден в течение всего времени просмотра фиксировать взгляд на дистанции расположения телевизора, а постоянно перемещает взор с близких объектов сюжета на дальние.

С другой стороны, свойство трёхмерной интерактивности даёт возможность телезрителю вести активный просмотр телепередачи. Включая режим интерактивности, он становится сам себе режиссером, который быстро осваивает технику наездов, панорамирования, смены ракурсов, то есть всех тех приёмов, которые, будучи ему неизвестными, механически удерживают его внимание, подчиняя воле телережиссера. Обладая профессиональным взглядом, человек может более критично относиться к телевизионной информации.

Но многие люди могут не захотеть включать режим интерактивности. Им нравится неосознанный сон наяву, который, как и обычный сон, не обязательно запоминается. Мы знаем, что внимание человека циклично: бодрствование – сон, урок – перемена. Мы усваиваем информацию, только если у нас есть время на её усвоение, то есть, имеется возможность повторить увиденное несколько раз в воображении. Непрерывный поток телевизионного видеоряда позволяет манипулировать нашим вниманием. Фильмы в жанре «экшен» могут запоминаться лишь в самых общих чертах, что и требуется для того, чтобы оставить память свободной, например, для рекламы.

Говоря о видеоряде, мы не должны забывать о *речи*. Обсуждая особенности захвата сознания, хотелось бы обратить внимание читателей на весьма странный способ общения людей между собой с её помощью. Произнося слово: «табуретка» я имею в виду сиденье с четырьмя ножкам, а мой собеседник столяр воображает целый комплекс технологических приёмов и свойств материала, необходимых для изготовления этого «сиденья»; тем не менее, благодаря контексту и снисхождению столяра мы понимаем друг друга. Такой способ кодирования визуальной информации с использованием личного опыта корреспондентов может легко приводить к недоразумениям, но заслуживает и пристального изучения, и не только в связи с возможностью гигантского сокращения с его помощью информационных потоков. Важно то, что речевое обращение к человеку на понятном ему языке переводит его в творческий режим, включая его воображение. Этот режим включается и при чтении человеком какого-либо текста. Таким образом, язык, слово способствуют рождению творца.

Однако фильмы с разговорами не слишком популярны у телезрителей, а информационные программы насыщены видеоинформацией и требуют минимального воображения. Наличие хорошего воображения само по себе ещё не обеспечивает творческого режима жизни. Необходима осознанность восприятия информации как внешней, так и порождаемой самим человеком.

Наркомания – принудительный, «химический» уход в мир сна, игромания – заикленное, азартное, выполнение автоматических движений, телемания – пассивное впитывание визуальной информации. Всё это формы неосознанности. Человек и в нормальных условиях обычно стремиться к неосознанным реакциям, переводя по мере обучения свои действия в разряд автоматических, потому что,

с одной стороны, такие действия быстрее и надёжнее, а с другой, не требуют напряжения. Но заложенный в нас на глубоком уровне *активный интерес* к жизни постоянно устремляет нас к новым впечатлениям и умениям. При этом овладеть умением можно только в состоянии осознанности.

В XXI информационном веке человечество сталкивается с новым вызовом цивилизации – сохранить развитую психическую сферу человека в условиях, когда стремление к впечатлениям легко удовлетворяется, а отсутствие разнообразных умений не влечет угрозу для жизни. Справиться с этим вызовом для него означает подняться на новую ступень своего развития. Мы уже сталкивались с подобным вызовом для нашего тела, когда физическая работа стала быстро уступать место умственной и нашли выход в спорте и физкультуре. Спорт мы связываем с соревнованием, физическую культуру с общим уровнем развития человека. В новых условиях внедрения объёмного телевидения, по-видимому, правильно будет говорить и о необходимости создания новых видов спорта, и о психической культуре социума.

Важным признаком осознанности является способность различать миры. Сознательный наблюдатель в любой момент времени понимает, находится ли он в телевизионном, физическом или воображаемом мире. Другим признаком осознанности является творческий режим человека, когда он, взаимодействуя с миром, воплощает свой замысел. В творческом режиме человек активно приобретает технологические знания – умения.

Различению миров и творчеству надо учить с детства. Мы учим малыша отличать сон от яви и строить песчаный замок. Осознающему себя и окружающие обстоятельства человеку объёмное телевидение не угрожает захватом сознания. Заметим, что трёхмерная интерактивность объёмного телевидения позволяет телезрителю в буквальном смысле иметь свою точку зрения. Вместе с тем, мы должны отдавать себе отчёт в том, что приобрести осознанность и избежать захвата сознания весьма не просто. В качестве спортивного тренинга осознанности могут применяться специальные телевизионные игры, выиграть в которые можно, только напрягая интеллект, а не используя быстроту реакции и эрудицию. По-видимому, нам предстоит сильно изменить культуру использования телевидения с учётом его влияния на человека на физиологическом и психологическом уровнях.

Поиск решений, позволяющих сохранить творческие силы человека в условиях развития телевидения, должен осуществляться творчески, методом проб и ошибок, с использованием принципа моделирования. При этом в нашем распоряжении: наука, как инструмент моделирования, государства как модели человечества, религии, как модели сознания.

Если содержание (контент) передач объёмного телевидения будет зависеть от его творцов, и мы можем только надеяться, что оно не разочарует нас, то техническая сторона объёмного телевидения гораздо более определённа и может быть здесь описана.

1.7. Три части телевизионной системы

Телевизионная система состоит из трёх основных частей: преобразователя «свет-сигнал» (телекамеры), преобразователя «сигнал-свет» (телевизора) и линии связи (радио или проводной) между ними. Такой вид имеет и простая телевизионная система передачи объёмного изображения. При этом к обычной цветной телекамере добавляется устройство определения пространственной формы снимаемых объектов, работающее с телекамерой в режиме совмещения растров. Способы определения формы могут быть различными, например, способ, разработанный автором [10] (1980), позволяет это делать с помощью импульсного оптического (инфракрасного) излучения. Устройство, работающее по этому способу, получило название – телевизионный лазерный локатор (ТВЛЛ).

Получаемый в результате четырехкомпонентный *RGBD*-видеосигнал *объёмной телекамеры* (рисунок 8), отличается от *RGB*-сигнала обычной телекамеры компонентой дальности *D*, модуляция яркости которой соответствует рельефу объекта наблюдения. Эту компоненту принято называть картой глубины (*depth map*), а *RGBD* объёмную телекамеру – *времяпролётной* (англ. *Time of Flight, ToF*), в соответствии с классификацией в [11]. Этот термин соответствует определению дальности до точек поверхности объектов наблюдения по времени запаздывания прихода к телекамере отражённого от них импульса подсветки.

Другой способ получения карты глубины связан с анализом диспаратности (различия) изображений стереопары стереотелевизионной камеры. Четырёхкомпонентный видеосигнал, поступающий на *объёмный монитор*, позволяет воссоздать в пространстве форму поверхности и цвет снимаемых объектов.

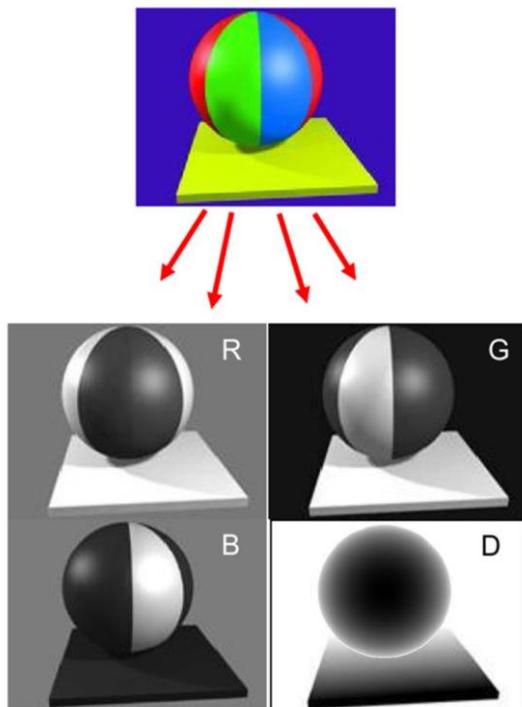


Рисунок 8 – Разложение изображения объекта на *RGBD*–компоненты

На рисунке 9 приведена функциональная схема такого монитора. Здесь красными линиями показана горизонтальная плоскость сечения поверхности изображаемого объекта, перемещающаяся равномерно вниз вдоль оси y со скоростью кадровой развёртки. Синей линией, лежащей в этой плоскости, показана траектория горизонтальной развёртки. Эта траектория формируется, как движением вдоль оси x лежащей на пересечении фронтальной и горизонтальной плоскостей, так и движением вдоль оси глубины z в соответствии с величиной D -компоненты видеосигнала. Цвет элемента объёмного изображения определяется остальными R

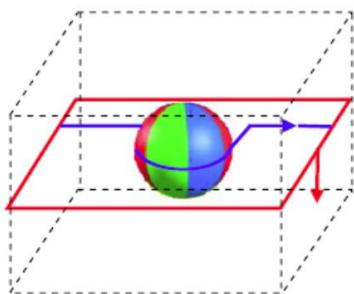


Рисунок 9 – Схема воспроизведения объёмного изображения объекта по *RGBD*-видеосигналу

(K – красной), G (Z – зеленой) и

В (С – синей) компонентами. Такая схема может использоваться в замкнутом мониторе (рисунок 3). При построении телевизионной системы по схеме рисунка 9, она, в соответствии с приведённой выше системой классификации (см. Таблицу 1), описывается формулой $6D(\theta, \varphi, r, x, y, z)$ или $6D(r, x, y, z)$, то есть является подлинно объёмной и допускающей перемещение зрителя в направлениях: z (ближе–дальше), x (влево–вправо) и y (вверх–вниз) без искажений изображения. Правда, эти перемещения существенно ограничены «гладкостью» изображения (отсутствием видимого пересечения краев объектов изображения при перемещениях зрителя).

Разработка объёмных (волюметрических, *volumetric*) мониторов имеет богатую историю [12] и основана на принципе воспроизведения объёмного изображения, состоящего из объёмных элементов – *вокселей* – материальных излучателей света, размещённых в рабочем объёме дисплея. В работе [13] описан объёмный дисплей на электронно-лучевой трубке (рисунок 10), где воксель формируется в точке вращающегося, покрытого фосфором стеклянного диска под действием электронного луча. Сканированием луча можно формировать воксель в любом месте сферы вращения диска. Это монохромный объёмный дисплей (1962 г.) очень низким разрешением и механическим приводом.

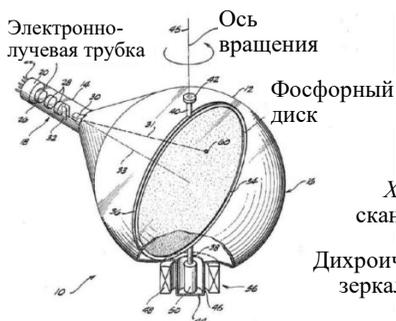


Рисунок 10 – Дисплей Р. Д. Кетчела на электронно-лучевой трубке

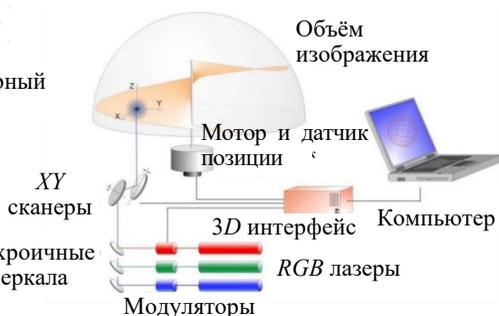


Рисунок 11 – Проект FELIX 3D. 10 000 вокселей с частотой кадров 20 Гц

Более совершенный дисплей [14] цветного изображения с лазерной подсветкой (2002 г.) содержит вращающийся экран (рисунок 11) в виде одного витка спирали, формируя воксели в цилиндрическом объёме. Сканирование пространства осуществляется зеркалами. Круговое оглядывание изображения приводит

к большому числу вокселей. Сложность расчётов положения вокселей на вращающемся теле и механическая развёртка приводят к низкому разрешению дисплея. Недостатком дисплеев с вращением экранов является полупрозрачность изображения. Много проблем также у дисплеев с транслюминесценцией и самосветящимися вокселями.

В объёмном мониторе, выполненном по технологии *DepthCube* (рисунок 12 а), объёмное изображение строится с использованием стопки экранов, которые прозрачны в обычном состоянии и последовательно активируются на короткое время для рассеивания света проектора. Проектор проецирует на активированный экран изображение соответствующего среза поверхности трёхмерного объекта (рисунок 12 б). Изображение такого объёмного монитора может обладать выраженной кулисностью, при недостаточном числе планов по глубине и тогда может возникнуть сильное ограничение на перемещение зрителя в направлениях: z , x и y из-за возможных искажений, вызванных наложением или разрывами планов изображения. Здесь телевизионная система также как в предыдущем случае имеет формулу $6D(r, x, y, z)$ с учётом пределов ограничений, а схема монитора соответствует замкнутому типу.

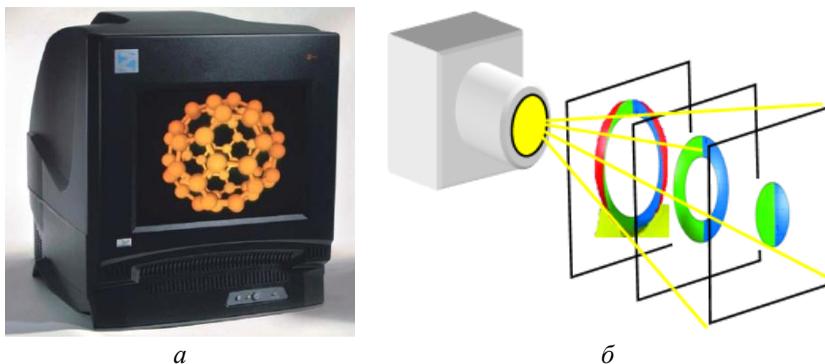
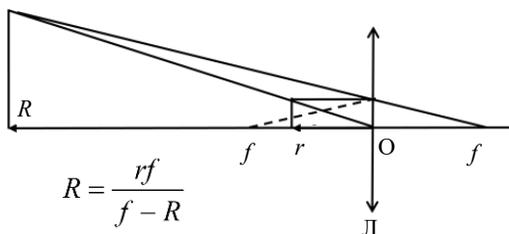


Рисунок 12 – Объёмный дисплей *DepthCube* с 20 планами глубины (а);
схема воспроизведения объёмного изображения объекта
с тремя планами глубины (б)

В принципе, схему каждого из этих мониторов можно изменить, перейдя к монитору открытого типа. Для этого перед окном монитора следует поставить большую оптическую линзу, работающую в режиме линзы. На рисунке 13 а приведена оптическая схема и формула такой линзы. Согласно формуле и схеме, при изменении глубины точки изображения замкнутой части объёмного

монитора z , находящейся в пределах фокусного расстояния линзы f , видимая (оптическая) глубина положения точки изменяется в широких пределах в соответствии с величиной R , эквивалентной физическому расстоянию этой точки в пространстве. Для обеспечения этой эквивалентности глубина изображения замкнутой части объёмного монитора должна быть нелинейной функцией глубины (дальности) расположения снимаемых объектов наблюдения. Другими словами, карта глубины должна подвергнуться нелинейному квантованию по глубине так, что плотность расположения близких слоев изображения будет выше, чем плотность далеких. Заметим, что такое требование соответствует разрешающей способности по глубине зрительной системы.

Любопытно, что установка линзы перед телевизором использовалась в первых образцах телевизоров для увеличения размеров изображения, как показано на рисунке 13 б. В нашем случае используется не способность линзы к увеличению, а способность к удалению изображения, т. е. свойство изменения кривизны фронта оптической волны света в зависимости от положения рассматриваемой точки в пределах фокусного расстояния линзы.



а
 б
 Рисунок 13 – Оптическая схема и формула линзы (а);
 телевизор КВН-49 с увеличительной линзой (б)

Угол *оглядывания* объекта наблюдения (смены ракурсов) при смещении головы телезрителя в мониторе с линзой невелик. Этот угол сравним с углом оглядывания объекта голограммы. Для плоской голограммы угол прямо пропорционален размеру окна голограммы и обратно пропорционален удалённости объекта от окна. В качестве окна в голографии может использоваться сама пластина голограммы. Для подчеркивания эффекта объёмности угол оглядывания является важным фактором, поэтому при голографировании объект съёмки стремятся расположить ближе к фотопластине.

Здесь и далее мы не рассматриваем голографический вариант системы объёмного телевидения поскольку голография требует от телевизионной аппаратуры очень высокого пространственного разрешения при фиксации интерференционной картины (голограммы), когерентных источников подсветки с большой длиной когерентности и тремя различными длинами волн для передачи цвета. При этом процесс голографирования чувствителен даже к малым вибрациям и сдвигам, а передача голографической информации требует очень широкой полосы частот канала связи. Большую технологическую проблему, по-видимому, также представляет создание «толстой фотоэмульсии» – трёхмерной среды регистрации считываемой голограммы, аналогичной оптической голограмме по методу Ю. Н. Денисюка. Следует отметить, что сам принцип голографии – полной записи (по терминологии Д. Габора) кодирует очень много ракурсов снимаемого объекта, что снижает требования к «гладкости» объекта при малом числе ракурсов съёмки.

Оглядывание объёмного изображения при движении головы зрителя – это лишь один из возможных вариантов. Другие варианты предполагают использование джойстика (компьютерной мыши, датчиков движения головы, нейроинтерфейса) для смещения (вращения) изображения в окне объёмного телевизора.

Повороты и смещения, а также создание иллюзии движения внутри объёмного изображения являются элементами трёхмерной интерактивности и возможны в тех случаях, когда имеется трёхмерная виртуальная модель физического мира – мир телевизионный. Наиболее удобный вариант объёмного монитора для движения внутри этого мира это наголовный дисплей (*Head Mounted Display*), используемый в системах виртуальной реальности. На рисунке 14 приведены фотографии наголовных дисплеев: наголовного дисплея (HMD) со стереокамерой (HMC), гибридным треккером для отслеживания направления взгляда и наушниками и ретинального дисплея с проекцией изображения непосредственно на сетчатку глаза, являющегося в данном случае элементом системы для создания дополненной (совмещенной) реальности, допускающей, в частности, оглядывание виртуального объекта наблюдения левым глазом с помощью джойстика, но без дополнительного создания объёмности. Дисплей на рисунке 14 *a* воспроизводит стереоскопическое изображение виртуальной среды, используя стереопару – два снимка среды с различных ракурсов, соответствующих ракурсам наблюдения двух глаз.



Рисунок 14 – Наголовный дисплей – а; ретинальный дисплей – б

Совмещение (слияние, фузия) кадров стереопары в сознании создает дополнительную иллюзию объёмности среды. Дополнительную потому, что кроме анализа диспаратности (различия) снимков стереопары мозг человека создаёт ощущение объёмности и по ряду других признаков объёмности, если объекты среды ему знакомы. К сожалению, метод стереоскопии, в его классическом варианте, содержит одну неприятную особенность – он не позволяет глазам динамично перефокусироваться (аккомодировать).

Плоские изображения стереопары всегда находятся на определённом постоянном расстоянии от наблюдателя, и, фокусируя глаза, человек вынужден, напрягать цилиарные мышцы глаз с одной и той же силой, как при чтении с плоского листа книги. Но, в отличии от режима чтения букв, где сведение глаз (конвергенция), разведение (дивергенция) и фокусировка (аккомодация) статичны и согласованы между собой, различно удалённые объекты сцены требует перестройки (вергенции) и аккомодации глаз для минимизации диспаратности и обеспечения чёткости изображения в зоне внимания (ясного зрения) человека. В природной среде при переносе взгляда на различно удалённые объекты происходит соответствующая удалённости дивергенция (конвергенция) глаз и соответствующая их аккомодация. Причём мозг определяет удалённость знакомых объектов ещё до выдачи команд на мышцы глаз. В виртуальной среде обычного стереоскопа возникают ситуации динамичной вергенции при переносе взгляда на различно удалённые объекты разглядываемой сцены, и отсутствия обеспечения при этом соответствующей аккомодации глаз. При попытке изменить аккомодацию изображение стереопары расфокусируется. Стереоскоп и окно (экран) стереоскопического, так называемого, 3D-телевизора вызывают у телезрителя повышенное утомление глаз и даже у некоторых людей болезненные ощущения, поскольку глаза телезрителя работают

в неестественном (не соответствующем опыту работы в природной среде) режиме. Это явилось причиной спада большого по началу спроса на объёмные стереотелевизоры. В таких телевизорах и мониторах используются следующие способы воспроизведения стереоскопического изображения: лентикулярный, эклипсный, поляризационный или даже анаглифический.

Лентикулярный способ предполагает наличие экрана, содержащего растр из длинных вертикальных цилиндрических линз малого радиуса. При правильном положении головы телезрителя, он видит на экране различные изображения стереопары разными глазами. Это безочковый вариант стереоскопического телевизора (монитора). В этом случае может создаваться несколько зон для положения головы в комнате для одновременного просмотра передачи несколькими телезрителями.

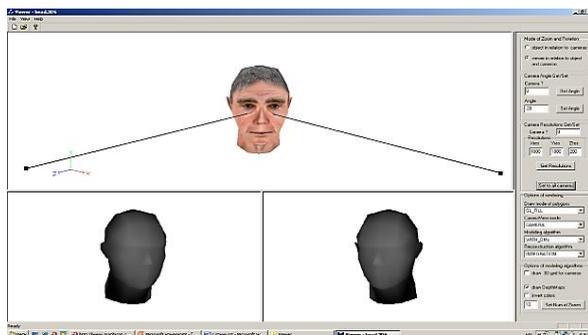
Эклипсный предполагает наличие специальных очков, стекла которых противофазно переключаются от прозрачности к полной непрозрачности синхронно с частотой телевизионных кадров (100 Гц). При этом на обычный экран проецируется поочередно изображения стереопары для левого и правого глаз. Число зрителей, смотрящих на этот экран ограничено лишь количеством комплектов эклипсных очков.

Поляризационный способ требует одновременной передачи на общий экран двух изображений стереопары с различной поляризацией излучаемого света и очков с различными стеклами – поляризационными фильтрами. Фильтры различаются направлением поляризации. При использовании линейной поляризации это скрещенные поляроиды; при круговой поляризации – поляризационные фильтры с различным направлением вращения. Использование круговой поляризации предпочтительнее.

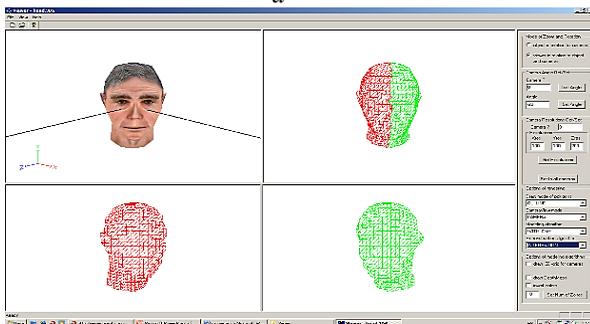
Анаглифический способ – самый простой, но даёт наихудшее качество цветопередачи. При этом стекла (пластик) очков выбираются разноцветными, например, левое синим, а правое красным. Левое изображение стереопары на общем экране окрашено в синих цветах, а правое в красных. То есть используется спектральная сепарация изображений стереопары. Оптимальным выбором спектральных характеристик стекол очков достигается не только сепарация, но и некоторое приближение к правильной цветопередаче, благодаря слиянию в сознании зрителя не только геометрических, но и цветовых составляющих стереопары.

Все перечисленные технические способы демонстрации стереоскопического изображения в настоящее время не обеспечивают согласованной вергенции и аккомодации глаз телезрителя. Однако, представляется возможным обеспечить такое согласование, благодаря некоторым особенностям зрения человека. Физиологические исследования [15] показали, что вергенция является более быстрым процессом, чем аккомодация. Вергенция возникает через 0,16...0,20 секунды после смены точки фиксации взора, а аккомодация через 0,4...0,5 секунды. Это обстоятельство можно учитывать для быстрого определения с помощью треккера угла вергенции глаз и использования его значения для управления параметром оптической системы телевизионного стереоскопа – изменения оптической силы её линз, обеспечивая правильную аккомодацию глаз. Такое техническое решение обеспечило бы комфортные (привычные) динамические условия работы глаз при рассматривании объёмных (трёхмерных) сюжетов, т. е. позволило бы избежать утомления глаз.

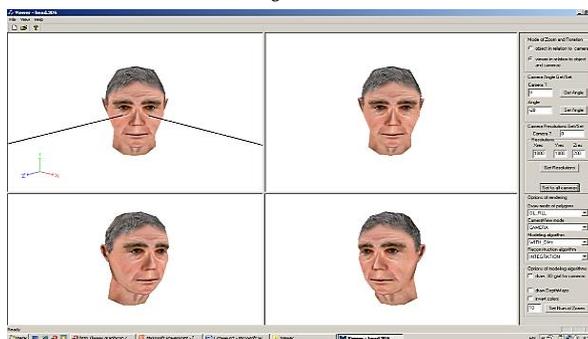
Обратимся теперь к вопросу связи между объёмной телекамерой и объёмным монитором. В объёмном телевидении их целесообразно связать через создание трёхмерной виртуальной модели физического мира – мира телевизионного. Эта целесообразность диктуется тем обстоятельством, что одна телекамера в каждый момент времени даёт только *один ракурс* наблюдения физического мира, а телезритель хочет иметь *произвольный ракурс* наблюдения в виртуальном мире, как в своем физическом мире и даже с большей возможностью доступа к объектам виртуального телевизионного мира. Очевидно, что для этого необходимо множество телекамер (ракурсов). При этом с экономической и эстетической точек зрения число этих камер следует минимизировать. В частности, две объёмные телекамеры, разнесенные на значительное расстояние, должны обеспечить решение задачи создания множества возможных ракурсов наблюдения. Результаты компьютерного моделирования «сшивки» двух ракурсов объёмных телекамер, выполненного нами в 2004 г., приведены на рисунке 15. Здесь мы можем видеть: 3D образ головы и карты глубины двух начальных ракурсов, каркасные модели ракурсов и их сшивка; наложение сшитой текстуры головы на каркасную сшивку и произвольные ракурсы головы.



a



б



в

Рисунок 15 – Компьютерное моделирование «шивки» двух ракурсов объёмных телекамер: 3D образ головы и карты глубины двух ракурсов – *a*; каркасные модели ракурсов (снизу) и их шивка (справа вверху) – *б*; наложение шитой текстуры головы на каркасную шивку (справа вверху) и произвольные ракурсы головы (снизу) – *в*

Моделированием было показано, что вид любого из множества выбранных ракурсов итоговой модели головы соответствует виду аналогичного ракурса её исходного 3D образа,

хотя итоговая модель получена всего из двух ракурсов. С помощью компьютерной мыши мы могли наблюдать любой ракурс модели в пределах угла между двумя начальными ракурсами.

Сшивка изображений от множества объёмных телекамер более сложная процедура, предусматривающая точное определение местоположения каждой из телекамер, ориентацию их оптических осей и значений их углов зрения. В любом случае важной задачей является минимизация количества необходимых для полной модели физического пространства телекамер. Удобным способом решения такой задачи может стать замена на начальном этапе телекамер на светильники с углами свечения, соответствующими углам зрения телекамер. Если мы хотим моделировать интерьер комнаты с мебелью (вещами), то должны разместить светильники так, чтобы исключить тени от предметов. Наличие теней, при замене светильников на телекамеры, приведёт к образованию на виртуальной модели комнаты «мёртвых зон» (разрывов модели) на месте теней, которые придется маскировать тем или иным способом. Хорошим результатом будет, если нам удастся осветить комнату без теней с помощью четырёх светильников, расположенных в углах комнаты. При съёмке комнаты для минимизации числа камер может оказаться полезным предварительно удалить из неё образующие тень предметы и создать отдельно их виртуальные образы – модели, которые будут помещены на виртуальное место предметов, соответствующее физическому месту, с которого они были временно удалены. Модели (аватары) также необходимы для подвижных объектов, например, людей. Такие модели могут строиться в процессе съёмки по мере их движения: перемещения и поворотов.

Телевидение, в отличие от факсимильной связи, предназначено для передачи движущихся изображений динамических сюжетов, то есть сюжетов, разворачивающихся во времени. Время в телевидении делится на фазы: накопления и считывания электрического сигнала. В фазу накопления осуществляется преобразование энергии света (фотонов) в электрический заряд, а в фазу считывания – фиксация уровня заряда и его обнуление. Разбиение на фазы создаёт дискретный характер изображения. Дискретами (фрагментами) изображения являются телевизионные кадры. Эти фрагменты могут как непосредственно передаваться телезрителю, так и запоминаться с последующим воспроизведением. Развитие технологий запоминания (консервации) движущихся изображений основано на том требовании, что частота телеви-

зионных кадров должна превышать предельную частоту мельканий зрительной системы. При этом видеоряд воспринимается как непрерывный. Кинематограф, родившийся раньше телевидения, показал, что частота кадров не должна быть меньше 48 Гц. Современные требования увеличивают эту частоту до 100 Гц. Запоминание кадров необходимо телевизионной системе не только для долговременной консервации, как при использовании видеоманитофонов, но и для эффективного кодирования телевизионного электрического сигнала. Важной задачей кодирования является сжатие данных телевизионной информации для снижения требований к пропускной способности системы связи [16], [17]. Одним из эффективных способов кодирования является передача межкадровой разности двух соседних кадров вместо второго кадра. При этом второй кадр восстанавливается в полном объёме при декодировании в точке приёма телевизионного изображения.

В общем случае телезритель объёмного телевидения (объёмного изображения и звука) при реализации свойств трёхмерной интерактивности и объёмности должен иметь возможность виртуального перемещения не только в пространстве, но и во времени. Возможность перемещения в пространстве–времени придаёт телезрителю новое качество. Он превращается из телезрителя в теленавта. Модель «окружающий мир – мир телевизионный» должна обеспечивать систему объёмного телевидения любого вида, который определяется описанной выше формулой, полным объёмом информации, фиксируемой пленоптической функцией $P = P(\theta, \varphi, \lambda, t, x, y, z)$, где помимо геометрических параметров представлены длина волны излучения элемента объекта λ и время t . При этом предполагается, что вместо спектра излучения фиксируется его цвет, определяемый параметрами: $R(K)$, $G(Z)$, $B(C)$, а время фиксируется с определённым интервалом дискретизации, как и все остальные параметры.

В телевидении термин *дискретизация* используется в основном в связи с пространством и временем. В связи с другими параметрам используется термин *квантование по уровню*. Важно отметить, что дискретизация (обобщенно) характеристик модели телевизионного мира не обязательно соответствует дискретизации модели используемой конкретной системой объёмного телевидения. Значения интервалов и числа уровней дискретизации на приёмном конце системы связи могут изменяться за счёт интерполяции, экстраполяции и прореживания значений величин в новых выбранных точках отсчёта. Такая практика имеет тем большее

значение, чем больше «интеллекта» у телевизора. В качестве примера системы связи с развитым интеллектом можно привести общение двух человек, где один (передатчик) *рассказывает* другому о месте, в котором побывал, а другой (приёмник) *воображает* это место. Говоря об интеллекте, то есть о способности к переработке информации, мы сталкиваемся с тем, что информация всегда представляется в дискретном виде, от поиска элементарных частиц материи до разделения знания на различные науки.

Специалисту в области цифрового сжатия телевизионной информации может показаться невероятной возможность сжатия и передачи по телевизионным каналам больших динамичных моделей окружающего мира. Между тем, следует принять во внимание, что существует несколько принципов редукции видеоинформации, эффективных именно для объёмного телевидения.

Во-первых, телезритель в каждый данный момент времени не видит всю модель, а лишь её часть, доступную его текущему восприятию, как если бы он воспринимал стереокино. Поддерживая двухстороннюю связь с сервером хранения модели, телезритель передает серверу необходимые координаты своего виртуального местоположения и угловые координаты направления оптической оси своего взора в данный момент времени: $(\theta_0, \varphi_0, t, x, y, z)$, – и сервер передаёт ему необходимую часть модели.

Во-вторых, обычно большинство объектов наблюдаемой сцены статичны (как на рисунке 16) и, следовательно, достаточно один раз загрузить в «умный» объёмный телевизор интерьер помещения, чтобы потом снимать его в любых ракурсах виртуальной стереокамерой без дозагрузки с сервера.



Рисунок 16 – Статическая часть телевизионной модели (интерьер музея)

Статичные объекты трёхмерной телевизионной модели, однократно введённые в память вашего объёмного телевизора, далее служат сценой для развёртывания динамичного сюжета передачи. Заготовки таких сцен очень больших размеров могут использоваться телережиссером для создания виртуальных декораций. Можно, например, один раз архивировать трёхмерный интерьер Эрмитажа и далее использовать его части в различных телепередачах. Для обычного телевизионного репортажа из музея с использованием технологии виртуальной телекамеры достаточно передавать телезрителю код положения и ориентации этой телекамеры, и он будет видеть предлагаемые части предварительно загруженной им в свой телевизор модели интерьера музея, слушая комментарий искусствоведа. В последнем случае информационный поток практически полностью определяется звуковым содержанием.

В-третьих, динамичные объекты сцены – персонажи (люди, животные) (рисунок 17) обладают ограниченным числом степеней свободы, то есть их возможное движение ограничено подвижностью суставов, числом и конфигурацией элементов их скелета.

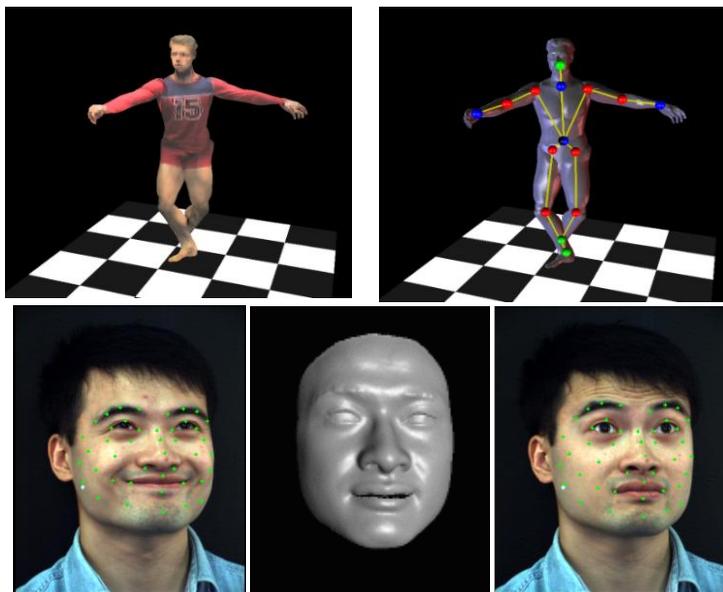


Рисунок 17 – Элементы динамической трёхмерной модели

На рисунке 17 точками обозначены узлы трёхмерных сеток (вершины), движения которых необходимо задавать. Анимиро-

ванная маска внизу в центре полностью воспроизводит мимику человека. Наложённые на сетки анимированные текстуры лица или тела, могут полностью имитировать их движения. Например, «скелет» автомобиля имеет всего пять элементов: кузов и четыре колеса. Задав всего два параметра: скорость вращения и поворот передних колес можно привести автомобиль в заданную точку статической модели. Естественные ограничения, накладываемые на траектории движения тел, приводят к модельному (*model-based*) способу кодирования динамичных объектов. Речь идет о создании отдельной модели каждого динамического объекта. Такая модель включает суставы (совокупности узловых точек движения), каркас (поверхность тела) и натянутую на каркас текстуру, задающую раскраску объекта. Даже относительно грубая динамическая модель позволяет кодировать реальную сцену с высокой степенью реалистичности, используя её для предсказания очередного пространственно-временного положения моделируемого объекта. Модельный способ позволяет ограничиться отслеживанием лишь параметров движения объектов, для чего могут использоваться даже простые, не объёмные телекамеры. Этот способ подразумевает возможность создания складов моделей одежды, обуви, предметов быта, моделей любых объектов естественного и искусственного происхождения. Единожды созданная и переданная модель обеспечивает создание громадного числа её ракурсов в пространстве по воле телережиссера или телезрителя.

В-четвертых, диапазон возможных яркостей объектов физического мира очень широк. Яркость в общем случае зависит от углов наблюдения θ и φ . Яркость хорошо рассеивающей поверхности одинакова в любом направлении (рассеяние по Ламберту). Яркость зеркальной поверхности соответствует яркости отражённых объектов и зависит от направления наблюдения. Во многих случаях мы имеем дело с промежуточной ситуацией, когда объекты не обладают зеркальной поверхностью, но яркость поверхности зависит от угла наблюдения. Перемещаясь в телевизионном пространстве (объёмном изображении) и наблюдая изменение ракурса объектов, мы должны видеть изменение их яркости, аналогичное изменению яркости объектов физического мира. Один из возможных подходов к кодированию угловой зависимости яркости заключается в её приближенном представлении с помощью сферических функций (рисунок 18).

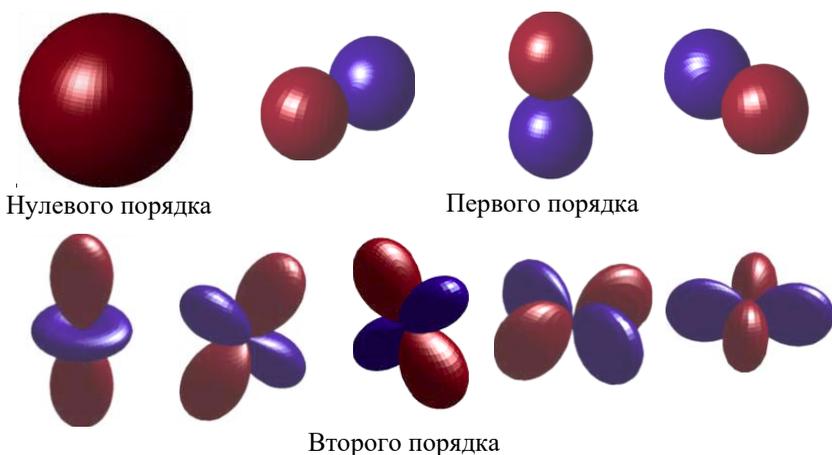


Рисунок 18 – Сферические функции (гармоники) трёх первых порядков

Такой подход подобен методу частотной фильтрации, используемой при обработке временных и пространственных сигналов, когда отбрасываются малые по величине высокочастотные составляющие сигналов. Частотное представление (преобразование Фурье) сигналов, в том числе многомерных, оказалось эффективным во многих областях. Например, в цифровом телевидении используется дискретно-косинусное преобразование для сжатия блоков двумерных изображений. Использование метода сферических функций возможно потому, что шероховатая поверхность объекта, в отличие от зеркальной, обладает фильтрующими свойствами, подавляя высшие гармоники функции углового распределения яркости точек поверхности. Кроме того, в случае сопоставления объёмного телевизионного изображения с представлениями телезрителя о физическом мире следует учитывать, что не многие люди способны проводить такое сопоставление с высокой точностью, тем более что обычно мы наблюдаем по телевизору незнакомые места и в условиях разнообразного освещения.

Рассеивающие свойства поверхности объектов наблюдения могут иметь выраженную спектральную и поляризационную зависимости, что позволяет в ряде случаев эффективно разделять зеркальную и рассеянную компоненты отражённого излучения. Зеркальная компонента излучения содержит информацию об отражённых объектах, которые сами содержатся в трёхмерной телевизионной модели. Существует возможность синтеза зеркаль-

ного отражения путём геометрических трансформаций соответствующих частей модели. Отмеченные обстоятельства позволяют говорить о перспективности использования сферических функций для целей сжатия больших телевизионных миров, поскольку во многих случаях можно ограничиваться сферическими компонентами первого и даже нулевого порядка.

В-пятых, когда из сцены выделены динамические объекты, мы можем предполагать, что движение многих из них подчинено законам физики, которые ограничивают возможные траектории их перемещений. Учёт законов движения позволяет использовать, например, кодирование с предсказанием движения объектов, что значительно снижает поток информации о движении.

Можно ли нарушать законы физики в телевизионном мире? Это должен решать, по-видимому, создатель телепередачи. Следуя своему художественному замыслу, он может позволить телезрителю, например, отключать гравитацию и тот, в интерактивном режиме, должен будет учиться двигаться в невесомости. С другой стороны, агрессивно настроенный телезритель может пожелать разрушить изнутри мир в своем телевизоре или нанести вред персонажам этого мира. Стоит ли поощрять подобные наклонности?

Поскольку трёхмерный телевизионный мир в значительной мере является копией физического мира, по-видимому, и в нём следует соблюдать многие физические и не только физические законы. Соответствующие ограничения, накладываемые на возможные трансформации телевизионных моделей, технически легко установить.

Перечисленные пять принципов снижения объёма передаваемой информации в процессе объёмного телевидения далеко не исчерпывают имеющиеся возможности сжатия информационного потока. Для поиска новых способов кодирования телевизионной информации и совершенствования известных необходимо познакомиться с особенностями кодирования зрительной информации в зрительной системе человека.

1.8. Зрительная система

Зрительная система человека обладает рядом интересных особенностей. Во-первых, разрешающая способность глаза быстро меняется при смещении изображения объекта от центральной ямки сетчатки к периферии (рисунок 19). Это обусловлено тем, что от

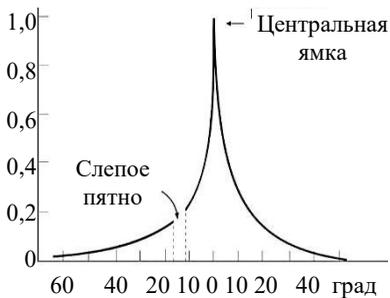


Рисунок 19 – Относительная разрешающая способность сетчатки глаза

сетчатки в мозг человека поступают не сигналы отдельных рецепторов (палочек и колбочек), а взвешенные суммы таких сигналов. Суммируются сигналы рецепторов небольших зон сетчатки – рецептивных полей. Рецептивные поля ганглиозных нейронов сетчатки имеют приблизительно круглую форму и оппонентную организацию. Диаметр рецептивных полей

изменяется от 2 угловых минут в центре до 1...2 градусов на периферии сетчатки. Для сравнения, угловой размер Луны составляет 30 минут, а колбочки 0,5 минуты. Рецептивные поля близких нейронов частично перекрываются. В результате такой обработки сигналов от 150 миллионов фоторецепторов сетчатки глаза в мозг идёт только миллион нервных волокон. При моделировании работы нейронов сетчатки в качестве весовых функций рецептивных полей используют двумерные функции Гаусса колоколообразной формы.

Во-вторых, каждый отдельный рецептор сетчатки имеет собственную биохимическую систему адаптации к световому потоку. Эта система может быть описана на функциональном уровне как система с глубокой параметрической обратной связью, аналогичная системе автоматической регулировки усиления (АРУ), используемой в радиотехнике. Учёт действия АРУ позволяет объяснить многие психофизиологические эффекты зрения, такие как наличие полос Маха (подчеркивание контуров) или появление последовательных образов (рисунок 20).



Рисунок 20 – Фиксируйте взгляд на центре флага 10 секунд, затем переведите его на белый лист бумаги (потолок). Вы увидите правильную раскраску флага

Параметры статического режима АРУ можно определить, например, используя психофизический закон Стивенса. Динамический режим АРУ возникает лишь тогда, когда входной сигнал значительно изменяется за время, сравнимое с постоянной времени цепи АРУ. В случае зрения человека эта постоянная примерно равна 1 секунде.

В-третьих, глаза человека постоянно двигаются. Повороты глаз: «влево – вправо», «вверх – вниз» и вокруг оптической оси производятся с помощью трёх пар мышц каждого глаза. С одной стороны, они позволяют стабилизировать поле зрения при наклонах и поворотах головы, а с другой отслеживать движущиеся объекты, сочетая плавные (нистагмы) и скачкообразные (саккады) движения. Особо следует отметить небольшие (менее 10 угловых минут) движения глаз, так называемые микросаккады, дрейф и тремор (рисунок 21), совершаемые бессознательно. Если искусственно остановить движения глаз, человек перестаёт видеть неподвижное изображение через несколько секунд, поскольку все участки изображения приходят за это время к одинаковому серому фону из-за перехода АРУ зрительной системы из динамического в статический режим.

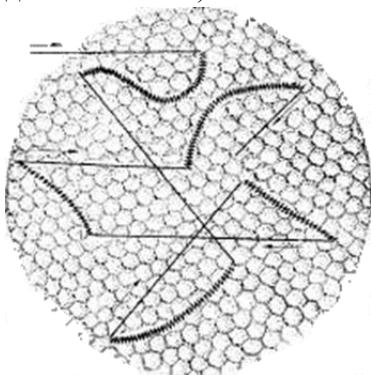


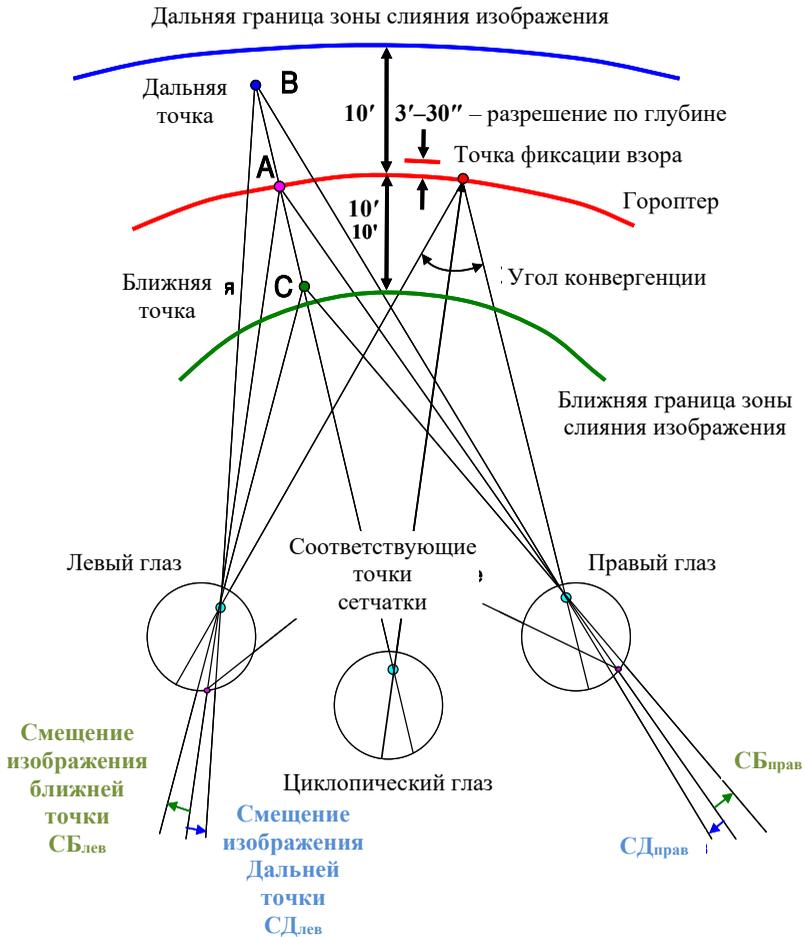
Рисунок 21 – Участок сетчатки глаза: микросаккады – прямые линии; дрейф – кривые линии; тремор – зигзаги на кривых

В-четвертых, бинокулярный (двухглазый) механизм зрения обеспечивает нам особое ощущение объёмности изображения трёхмерных объектов. То есть даёт нам

возможность определять относительную глубину точек поверхности объекта наблюдения.

Абсолютную глубину или дальность светящейся точки в физике часто соотносят с её параллаксом – углом, под которым из этой точки виден базис наблюдения. Базисом зрительной системы является расстояние между зрачками (в среднем 65 мм). При наблюдении светящейся точки из двух точек пространства сначала определяют её параллакс, а затем по параллаксу вычисляют расстояние до неё. Такой способ определения дистанции, оправданный геометрией наблюдения, закрепился в астрономии в названии единицы длины – парсек (параллакс – секунда),

соответствующей дистанции с которой базис (диаметр орбиты Земли) виден под углом в одну угловую секунду. Когда мы фиксируем взор на какой-либо точке, наши глаза сводятся на угол конвергенции (рисунок 22), который равен параллаксу этой точки.



$$|СД_{лев}| + |СД_{прав}| = \text{диспаратность дальней точки (+)}$$

$$|СБ_{лев}| + |СБ_{прав}| = \text{диспаратность ближней точки (-)}$$

Рисунок 22 – Бинокулярная диспаратность – основа стереоскопического зрения

Все другие точки пространства, которые *кажутся* нам находящимися на одинаковом с точкой фиксации расстоянии, находятся на некоторой поверхности, называемой гороптером. Гороптер можно определить опытным путём. Его форма зависит от расстояния до наблюдаемой точки. Хотя мы смотрим двумя глазами, но обычно видим один не двоящийся мир. Мы как бы видим его одним циклопическим глазом. Волшебство этого глаза в том, что он даёт нам одновременно ощущение объёмности и слитности изображения. Его срединное положение (рисунок 22) может меняться при амблиопии, смещаясь в сторону ведущего глаза. Задав положение циклопического глаза и его линию визирования выбранной точки пространства, можно ввести понятие диспаратности для пары изображений левого и правого глаза этой точки. Диспаратность равна сумме модулей отклонений этих изображений от изображений точки гороптера, пересекемой линией визирования циклопического глаза. Для точки пространства дальше гороптера диспаратности присваивается положительное значение, а для точки ближе гороптера – отрицательное. Можно показать, что диспаратность точки B (C) на рисунке 22 равна разности параллаксов точек A и B (C), то есть так называемому дифференциальному параллаксу.

Человек способен оценивать диспаратность точек локальных особенностей поверхности объекта наблюдения с помощью специальных бинокулярных нейронов зрительной коры головного мозга. Каждый такой нейрон работает, получая одновременно сигналы рецептивных полей левого и правого глаза. Весовые функции этих полей часто представляют действительной и мнимой частями комплексного вейвлета Габора. Эти функции не обладают круговой симметрией, в отличие от весовых функций первичных рецептивных полей ганглиозных нейронов сетчатки, но на уровне коры мозга может происходить синтез из первичных других рецептивных полей с необходимыми функциями.

Здесь мы предложим модель определения диспаратности с экспоненциальными весовыми функциями рецептивных полей левого и правого глаза. На рисунке 23 вверху слева показан график поверхности такой функции. Для наглядности поверхность модулирована по яркости в соответствии с высотой её точек. Справа вид графика сверху, дополненный линиями равного уровня. Снизу изображена модель, на которой кроме рецептивных полей левого и правого глаза показаны элементы логарифмирования и вычитания

входных сигналов. Выходы модели соответствуют двум отдельным бинокулярным нейронам. Один нейрон (+ Δ) реагирует на положительную разность входных сигналов, которая пропорциональна положительной диспаратности точек изображений светлого объекта (линии) на рецептивных полях. Другой нейрон ($-\Delta$) реагирует на отрицательную разность входных сигналов, если диспаратность светлого объекта отрицательна.

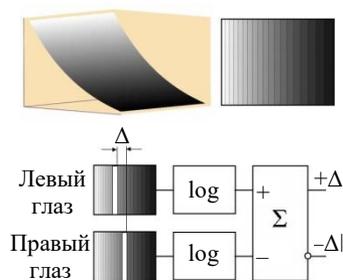


Рисунок 23 – Модель определения диспаратности светлого объекта

Совокупность реакций соответствующих бинокулярных нейронов даёт карту положительных или отрицательных диспаратностей по полю зрения, которая аналогична картам глубины, приведённым выше, но отличается от них фрагментарностью, соответствующей местоположению в поле зрения локальных особенностей яркости. Модель не реагирует на диспаратность изображений точек однородно окрашенной поверхности, что соответствует свойству нашего зрения. Анализ работы модели при различных оптических сигналах показывает, что минимизация ошибок определения диспаратности достигается при размере локальных неоднородностей яркости окраски объекта много меньше размера рецептивных полей. Кроме того, угловой интервал между неоднородностями должен быть много меньше углового значения диспаратностей. Эффективным средством минимизации ошибок является выделение контуров неоднородностей. Именно такую операцию производит совместное действие АРУ зрительной системы и механизма микросаккад. Обратим внимание на то, что максимальная амплитуда микросаккад (рисунок 21) одного порядка с максимальной глубиной зоны слияния (рисунок 22) – 10 угловых минут. Последнее обстоятельство, наряду с конфигурацией траекторий микросаккад и дрейфа, позволяет эффективно усреднять карты диспаратностей по времени.

На рисунке 23 приведена схема модели, реагирующей на повышенную яркость или локальное включение света. Модель, реагирующая на пониженную яркость (темный объект) или локальное выключение света отличается от нее наличием инверторов по входам сигналов.

Авторы [18] Поггио и Фишер обнаружили четыре типа бинокулярных нейронов с модуляционными характеристиками, приведёнными на рисунке 24. Наша модель объясняет возбуждение нейрона на дальнюю и ближнюю диспаратности. Модель, получающаяся суммированием выходных сигналов $+\Delta$ и $-\Delta$, объясняет торможение нейрона на нулевую диспаратность, а эта же модель с инверсией выходного сигнала объясняет возбуждение нейрона на нулевую диспаратность.

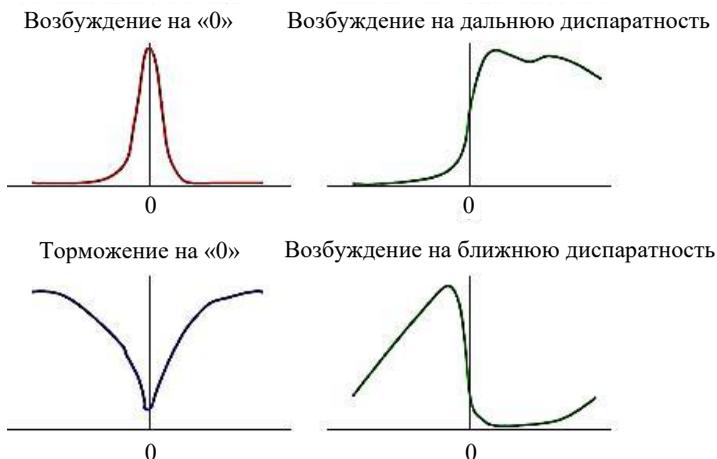


Рисунок 24 – Модуляционные характеристики бинокулярных нейронов

Особо отметим, что приведённая модель обладают инвариантностью по отношению к изменениям амплитуды и местоположения оптических сигналов. Так как вычисляемая разность логарифмов двух величин соответствует отношению этих величин, то одинаковые коэффициенты яркости и сдвига оптических сигналов сокращаются. Коэффициент разности сдвига экспоненциального вида при логарифмировании даёт линейную зависимость выходных сигналов модели от диспаратности.

Бинокулярное зрение человека, обеспечивающее фузию (слияние) изображений, формируемых мозгом с помощью каждого из двух его глаз, представляет собой весьма сложный процесс, требующий изучения. Для проектирования системы объёмного телевидения важно то, что глубина видимого пространства может быть разделена на три зоны. Наблюдение до трёх метров, сопровождается изменением, как конвергенции, так и аккомодации. От трёх до шести метров изменяется только конвергенция.

До 1300...2600 метров мы способны видеть рельефность (глубину) сцены, благодаря оценке диспаратности. Разрешение по глубине (диспаратности) зрительной системы человека варьируется от 3 до 30 угловых секунд у разных людей. Эти данные позволяют рассчитать число различимых человеком планов по глубине. Расчёт показывает, что на расстояниях от 1 метра до «бесконечности» число планов равно 450 при разрешении по глубине в 30 угловых секунд.

Отметим, что, например, конструкция объёмного монитора, приведенная на рисунке 9, обеспечивает возможность создания сотен и даже тысяч планов разрешения по глубине, а реализованная конструкция монитора, приведенная на рисунке 12 обеспечивает 20 планов разрешения по глубине, при этом разработчики утверждают, что им удалось сгладить дискретность (кулидность) изображения, обусловленную низким разрешением по глубине, с помощью оригинального алгоритма сглаживания.

Важными являются пороговые характеристики зрения. Порог конвергенции оценивается смещением объекта в глубину на 3 угловые минуты; порог аккомодации – на 30 угловых минут. При превышении порога зрительная система автоматически перестраивается на новую глубину.

В школьных учебниках по биологии и даже в учебниках по телевидению глаз иногда уподобляли фотоаппарату или телекамере. Как мы имели возможность убедиться – это весьма упрощённое представление. При разработке системы объёмного телевидения, когда перед нами встает проблема создания трёхмерной, динамичной модели (карты) физического мира, по существу, только зрительная система человека служит нам примером её решения. Пример может показаться трудно достигаемым даже на первых (низких) уровнях зрительной системы. Вызывает восхищение её эффективность при решении задачи адаптации.

1.9. Адаптация

Мы живем в мире, наполненном светом. Диапазон изменения яркостей окружающих объектов может достигать пятнадцати порядков (квадратично раз). Это диапазон изменения значений вышеупомянутой пленоптической функции. На рисунке 25 приведён диапазон яркостей белой бумаги, освещаемой различными источниками освещения.

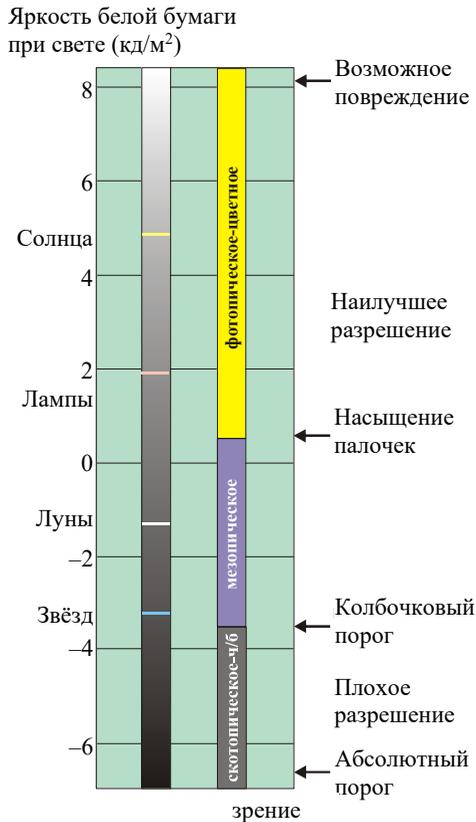


Рисунок 25 – Диапазон яркостей белой бумаги при различных источниках освещения

Даже создание такого диапазона яркостей в экспериментальных условиях представляет известную проблему. Во всем этом диапазоне глаз сохраняет работоспособность, хотя как жаловался Ньютон (1691 г.) в своем письме Локку, ему пришлось троё суток сидеть в полной темноте, чтобы избавиться от послепредобразов Солнца после наблюдений его в телескоп. Бойль (1664 г.) в таких опытах получил ожог сетчатки и продолжал видеть образ Солнца с закрытыми глазами даже спустя 10 лет. С другой стороны, после нескольких часов, проведённых в темноте, человек может видеть ничтожную вспышку света, эквивалентную восприятию десяти фотонов. Ни одна телевизионная камера не может перекрыть такой диапазон яркостей. В лучшем случае мы можем попытаться перекрыть его с помощью трёх телекамер разных диапазонов,

охваченных глубокими параметрическими обратными связями автоматических регулировок режимов (коэффициентов усиления).

При использовании трёхмерной интерактивности объёмного телевидения можно представить себе следующие виртуальные ситуации: сначала Вы находились в помещении, а затем вышли из дома полюбоваться закатом или, двигаясь в автомобиле по ярко освещенной автостраде, заехали в тоннель. Перепады яркости, вызванные таким Вашим поведением, весьма велики и для реалистичности объёмного изображения телевизионная модель должна содержать информацию о текущей освещённости на местности.

Актуальной является проблема отображения тона (ТМ – *Tone Mapping*) – сжатия широкого динамического светового диапазона (HDR – *High Dynamic Range*) в визуализируемый узкий динамический диапазон (LDR – *Low Dynamic Range*). Например, в Минесотском Институте прикладной математики был разработан (2009) алгоритм ТМ на основе модифицированного уравнения Нака–Раштона, позволяющий не только сжимать HDR, но и осуществлять локальное повышение контраста изображения и устранять потери цвета.

Необходимо отметить, что разброс коэффициентов отражения поверхностей природных объектов невелик, обычно не более 40 раз. Примерно такой же диапазон яркости отображает зрение человека без адаптации. Классический пример отображения с адаптацией связан с чтением книги. В помещении текст нам кажется черным, а лист книги белым. На ярком солнечном свете текст, все равно нам кажется черным, а лист белым, хотя объективно яркость текста на солнце в десятки раз больше яркости листа в помещении. Учитывая такие адаптационные свойства зрения достаточно иметь монитор с тысячекратным изменением яркости для воспроизведения любых изображений, в том числе и с имитацией пересветки. Современные цифровые фотокамеры обладают большим динамическим диапазоном и фиксируют HDR в RAW файлах (до 14 бит на канал). Хранение данных можно осуществлять в формате Tiff (англ. *Tagged Image File Format*) файлах (до 64 бит на канал). Для сравнения в широко используемом стандарте *True color* яркость (цвет) кодируется всего 256 уровнями (8 бит на канал), при этом глаз не ощущает «ступенек» тона. Обладая запасом разрешения по градациям яркости можно растягивать при воспроизведении диапазон яркости малоконтрастных локальных участков, повышая их локальный контраст.

Благодаря процессу эволюции, подвергаясь непрерывным колебаниям освещённости при смене дня и ночи, мы получили механизм зрения, обладающий феноменальными динамическими характеристиками. Обычно говорят об адаптационной роли зрачка глаза (изменение площади в 16 раз) и подчеркивают наличие двух видов фоторецепторов сетчатки глаза: палочек и колбочек, которые делят диапазон воспринимаемых яркостей на два. В действительности основная «тяжесть» адаптационной перестройки падает на биохимический комплекс механизма трансдукции фоторецепторов (изменение передачи в миллионы раз), а также на нейросетевой механизм организации рецептивных полей сетчатки глаза. Механизм фототрансдукции [19] обеспечивается длинным

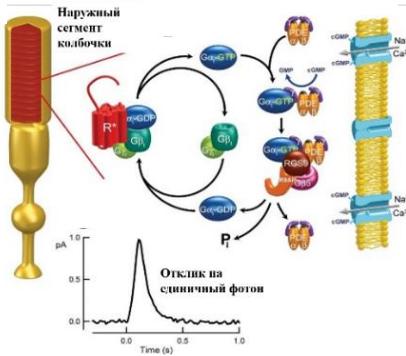


Рисунок 26 – Механизм фототрансдукции рецептора сетчатки глаза

каскадом биохимических реакций (рисунок 26), заканчивающийся цГМФ (циклический гуанин моноосфат) регуляцией пропускания мембранных каналов рецепторов. Изменение ионной проводимости мембраны фоторецептора вызывает изменение её внутриклеточного потенциала (минус 35 мВ). Увеличение яркости света вызывает гиперполяризацию мембраны, до минус 70 мВ, уменьшение – деполяризацию, до 0 мВ. Градуальные

потенциалы рецепторов воздействуют на биполярные и горизонтальные нейроны сетчатки глаза и далее на её амакриновые и ганглиозные нейроны так, что у ганглиозной клетки образуется приблизительно круглое рецептивное поле, с выраженными центром и периферией.

Упрощённая схема восприятия света для дневного (колбочкового) зрения показана на рисунке 27 (а, б, в). Три вида колбочек сетчатки глаза, изображены разными цветами в соответствии с их спектральными характеристиками: *L* (длинноволновые) – красным, *M* (средневолновые) – зеленым и *S* (коротковолновые) – синим. Каждая из колбочек обладает собственной системой автоматической регуляции усиления (АРУ).

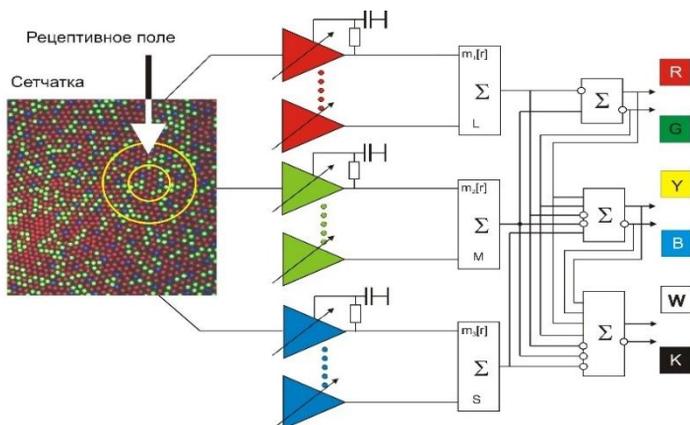
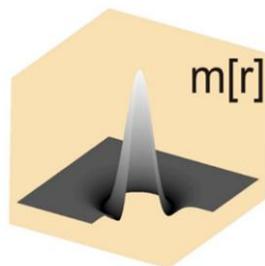


Рисунок 27 а – Функциональная схема восприятия света первичного уровня зрительной системы

Рисунок 27 б – Вейвлет «сомbrero» (с вырезанной для наглядности четвертью) представляет собой весовые функции трёх сумматоров сигналов колбочек концентрических рецептивных полей (ПП) нейронов R , G , Y , B , W и K



Свет	Образ на рецептивном поле при максимальном сигнале на выходе нейрона					
Вкл. (on)	R	G	Y	B	W	K
Выкл. (off)	G	R	B	Y	K	W
Образ						

Рисунок 27 в – Таблица реакций нейронов на различные образы

На схеме рисунка 27 а работа колбочек показана с помощью усилителей, изображенных треугольниками, охваченных цепями параметрической обратной связи, регулирующей коэффициент их усиления. Запаздывающий характер этой связи изображён с помощью интегрирующих RC -цепочек. Постоянная времени АРУ колбочек составляет величину порядка нескольких секунд. Роль обратной связи в механизме фототрансдукции, как считают, выполняют в основном ионы кальция, воздействуя на различные её

фазы. АРУ колбочки обладает статической и динамической характеристиками, последняя из которых соответствует периодическим скачкам взгляда с яркого на тёмный фрагмент объекта наблюдения. Важной особенностью усилителя с АРУ является различие его коэффициентов усиления для нижних и верхних *временных* частот, что соответствует различным наклонам его статической и динамической характеристик. Сигналы колбочек каждого вида с площади, очерченной большим кругом рецептивного поля, поступают на соответствующий сумматор. Каждый сумматор осуществляет взвешенное суммирование сигналов колбочек. Весовые функции сумматоров, можно аппроксимировать двумерными вейвлетами типа «сомбреро» (*mexhat*). Здесь положительная (возбудительная) часть вейвлета соответствует внутреннему (малому) кругу рецептивного поля, а отрицательная (тормозная) его внешнему кольцу. Центр слегка преобладает над периферией в отличие от принятого в вейвлет анализе баланса, что позволяет фиксировать постоянную составляющую сигнала. Выходные сигналы сумматоров обрабатываются согласно приведённой схеме. Небольшие белые кружки на входах и выходах элементов схемы обозначают инверсию сигналов. Результат обработки представляется сигналами на шести выходах: *R* (*red* – красный); *G* (*green* – зеленый); *Y* (*yellow* – желтый); *B* (*blow* – синий); *W* (*wait* – белый) и *K* (*black* – черный). Эти выходы соответствуют шести отдельным нейронам, с одинаковыми по площади, местоположению и форме рецептивными полями. Выходные сигналы здесь положительные (соответствуют деполяризации нейрона) и могут присутствовать либо на выходе *R*, либо на выходе *G*, либо отсутствовать на обоих. Аналогичное ограничение распространяется на пары *Y*, *B* и *W*, *K*. Такую организацию распределения выходных сигналов принято называть оппонентной. Согласно таблице на рисунке 27 в при проецировании на рецептивное поле подходящего образа, при включении или выключении света будет получена максимальная реакция (сигнал) соответствующего нейрона. При освещении рецептивного поля однородным по цвету излучением реагирует нейрон, соответствующий цвету излучения. Важной особенностью работы данной схемы является небольшой коэффициент усиления для сигналов нижних *пространственных* частот и большой для частот, на которые настроен вейвлет. Необходимо отметить, что разбалансировка вейвлета может меняться при увеличении внешней

освещённости в сторону уменьшения относительного веса его положительной части, то есть в зрительной системе присутствует ещё одна петля обратной связи АРУ, не показанная на схеме.

Приведённая схема имеет иллюстративное значение (показывает наличие различных механизмов зрения) и не может приниматься за вполне адекватную модель обработки сигналов зрительной системой человека. Элементы более точной модели можно найти, например, [20], где учитывается действие колбочек, биполярных и горизонтальных клеток сетчатки глаза. В частности, там учтён нелинейный характер передачи сигналов в сетчатке. Наша схема соответствует оппонентной теории цветовосприятия, предложенной Эвальдом Герингом в её линейном приближении. Учёт нелинейности можно свести к утверждению о том, что оппонентные выходы представляют не разность, а отношение входных световых сигналов из-за логарифмирования этих сигналов фоторецепторами.

Мы ощущаем красный цвет, когда возбуждаются нейроны R , при более активном действии колбочек L -типа, или ощущаем зеленый цвет, когда возбуждаются нейроны G , при преобладании действия колбочек M -типа, либо не ощущаем этих цветов при балансе действия этих типов колбочек, когда ни R , ни G не возбуждаются.

Мы ощущаем желтый цвет, когда возбуждаются нейроны Y , при преобладании суммарного действия колбочек L и M типов над действием колбочек S типа, или ощущаем синий цвет, когда возбуждаются нейроны B , при преобладании действия колбочек S типа, либо не ощущаем этих цветов при балансе действия этих типов колбочек. Обратите внимание на то, что по схеме суммарное действие колбочек L и M типов берётся без значения разности их действия.

Мы ощущаем степень светлоты объекта наблюдения, когда срабатывают нейроны W или степень его затенённости, когда срабатывают нейроны K по сравнению со средним уровнем ощущения яркости сцены.

Градуальные выходные сигналы схемы в реальной сетчатке преобразуются ганглиозными клетками в частотно-импульсные сигналы потенциалов действия и по их аксонам, которые составляют зрительный нерв, передаются в головной мозг.

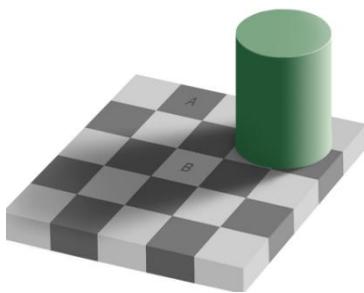


Рисунок 28 – Тест Эдельсона

с технической точки зрения, поскольку речь идёт о возможности резкого сокращения динамического диапазона светоприёмника.

Как отмечалось раньше, весовые функции рецептивных полей сетчатки увеличивают свой масштаб по мере удаления от её центра, что приводит к снижению разрешающей способности сетчатки на периферии (см. рисунок 19). Это соответствует возможности вычисления вейвлет-спектра изображения по осям сдвига и масштабирования, накапливая данные при разнообразных скачках (саккадах) глаз.

С другой стороны, во время дрейфа глаз (см. рисунок 21) изображение с небольшой скоростью, до 20 элементов в секунду смещается по сетчатке глаза. В телевидении используются телевизионные камеры ИК – диапазона, в которых реализовано так называемое оптическое микросканирование (искусственный дрейф) с помощью вращающейся наклонной стеклянной плоскопараллельной пластинки [21]. Благодаря этому техническому приёму разрешающая способность телекамер существенно повышается. Вероятно, дрейф глаз используется зрительной системой человека на её высоком уровне с той же целью, частично преодолевая потерю разрешения в центральной области поля зрения.

Мы рассмотрели ряд механизмов начальных стадий работы зрительной системы человека. Уже на этих стадиях нас поражают её возможности парировать гигантские изменения освещённости окружающей среды и добиваться высокого пространственного и временного разрешения при наличии весьма скромных возможностей по быстрдействию.

Эволюция биологических и технических систем отображения окружающего мира идёт, на первый взгляд, различными путями. Между ними имеется, однако, много общего, начиная с подобия

Снижение удельного веса нижних пространственных частот приводит к эффекту, хорошо видному на тесте Э. Эдельсона (рисунок 28), где клетки доски А и В, кажутся различными по яркости, хотя на самом деле таковыми не являются.

Организация преобразователя «свет-сигнал», которую нам демонстрирует зрительная система человека, заслуживает пристального внимания

генных мутаций и изменчивости технических идей и заканчивая прямым заимствованием молодым эволюционным потоком технических решений у старушки природы.

Сигналы, поступающие в мозг от сетчаток глаз, подвергаются им сложной обработке. На корковом уровне работы зрительной системы производится выделение направлений контуров и движений объектов наблюдения с помощью сложных рецептивных полей. С помощью метода функциональной магниторезонансной томографии удалось выявить у человека до 36 зон коры мозга, осуществляющих последовательную обработку зрительной информации. По-видимому, обработка идет по принципу детектирования все более обобщённых признаков распознаваемых объектов наблюдения. Обнаружены механизмы синхронизации процессов обработки в виде волн электрической активности мозга. Изучая мозг, можно видеть, что процесс анализа информации о внешнем мире идет параллельно с её сжатием.

1.10. Трансформации мира телевизионного

Если создана трёхмерная модель физического мира, возможны разнообразные преобразования или искажения такой модели. Можно сжимать, растягивать, наклонять, перемещать, переворачивать и подвергать различным нелинейным искажениям, единожды созданный телевизионный мир. При этом, кроме геометрических, легко доступны цветовые и тональные трансформации. Формулы трансформаций могут включать самые разнообразные математические операции.

Например, рисунок 29 иллюстрирует операцию склейки тессеракта – гиперповерхности четырехмерного куба. Представим себе, что мы находимся в одном из желтых кубов. Мы можем пойти в каком-либо направлении и пройти сквозь грань этого куба, оказавшись в соседнем. Если грань наружная, то в соответствии со стрелкой «склейки» граней мы все равно окажемся в другом соответствующем кубе. В каком бы кубе мы первоначально ни находились, идя по прямой и пройдя три промежуточных куба, мы снова окажемся в том, из которого вышли. Аналогичное, но двумерное, путешествие предпримет муравей, переползая с грани на грань обычного детского кубика. Операция склейки часто используется математиками при топологических преобразованиях. В нашем случае внутренностью различных кубов могут быть

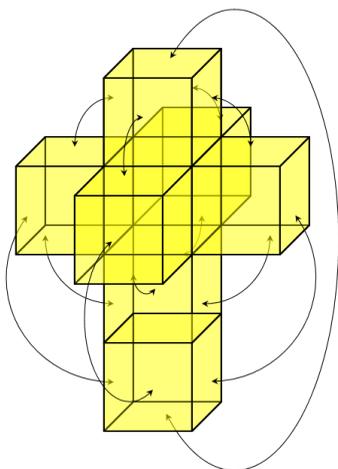


Рисунок 29 – Трёхмерная развертка четырёхмерного куба; стрелки показывают какие плоскости следует «склеить»

интерьеры различных комнат, и мы будем испытывать иллюзию прохождения сквозь стены этих комнат, каждый раз оказываясь в новой и блуждая по кругу как в лабиринте. Можно наполнить комнаты зеркальными по отношению друг к другу интерьерами, где левое меняется на правое. Вместо комнат могут быть более обширные области пространства. Вместо кубов могут быть другие многогранники. Можно представить себе путешествие по трёхмерным гиперплоскостям более чем четырёхмерных многообразий. Склейки можно осуществлять не только в пространстве, но и во времени.

До сих пор мы ограничивались обсуждением визуальной (световой) части объёмного телевидения. Объёмный звук, по-видимому, будет неотъемлемой частью телевизионного мира.

1.11. Объёмный звук

При восприятии физического мира (см. 1 на рисунке 6), мы пользуемся различными органами чувств (модальностей): зрения, слуха, осязания, обоняния и вкуса. Можно говорить и об органе прошедшего времени – памяти и об органе будущего времени – сознании. Причём последний позволяет предсказывать будущее на основании данных предыдущих органов.

При восприятии виртуального телевизионного мира (2) для полной иллюзии реальности необходимо иметь возможность также задействовать все наши чувства. Восприятие, т. е. построение воображаемого мира (3) в нашем сознании подобно совмещению в одной модели нескольких чувственных миров: *зрительного, слухового, осязательного, обонятельного и вкусового*. При этом телевизионный мир должен совмещать в себе соответствующие миры: *световой (изображения), звуковой, тактильный, и химический (газовый и жидкостный)*.

Идея вложенности (совмещения) различных миров популярна как у мистиков, так и у физиков-теоретиков с их моделью мультивселенной. С практической информационной точки зрения эта идея полезна при решении проблем совместимости и синхронизации миров. Причём число выделенных миров может превышать общепринятое число органов чувств. Например, мы можем разделить световой (зрительный) мир на красный, зелёный и синий и решать проблему совмещения цветов. Такая проблема была весьма актуальна в эпоху телевизионных цветных кинескопов. Если мы стремимся создать имитацию тактильных ощущений, например, с помощью перчаток с встроенными ультразвуковыми вибраторами, то эти ощущения должны быть согласованы со световой виртуальной моделью в процессе её «ощупывания», т. е. следует решить задачу совмещения светового и тактильного миров.

Звуковой мир, как составная часть физического, соединяется с воображаемым миром с помощью системы передачи звука, состоящей из трёх основных частей: преобразователя «звук-сигнал» (микрофона), преобразователя «сигнал-звук» (громкоговорителя) и линии сети связи. Такой вид имеет и простая звуковая система передачи объёмного звука, содержащая два микрофона и два громкоговорителя, схема приёмной части которой приведена на рисунке 30. Такую систему можно определить формулой, как и в случае изображения $4D(\varphi, x, y, z)$, согласно которой слушатель с координатами в пространстве: x, y, z может ощущать направление на источник звука в пределах диапазона угла φ , равного, например, 60 градусам. Виртуальные источники звука могут находиться на линии, соединяющей два громкоговорителя. Такую возможную локализацию виртуальных источников звука трудно назвать удовлетворительной.

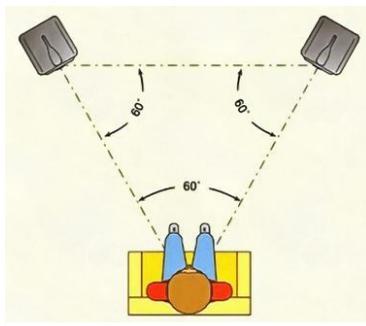


Рисунок 30 – Звуковая стереосистема

В последние годы в системах виртуальной реальности наметилась тенденция к предоставлению пользователю 360-ти градусного звука. Самый популярный аудиостандарт для обработки и доставки такого звука называется *Ambisonics* [22]. Основной подход этого стандарта в том, чтобы рассматривать звуковую сцену как полную 360 – градусную сферу звука, приходящего с разных (любых) направлений (перифония) в центральную точку, где

находится особый микрофон, типа *SoundField*, содержащий четыре капсуля с кардиоидными характеристиками, во время записи или точка наилучшего восприятия слушателя во время воспроизведения звука с помощью выбранного набора громкоговорителей.

На рисунке 31 приведен внешний вид микрофона RODE *SoundField* NT-SF1. Исходные сигналы капсулей микрофона формата *A* преобразуются в соответствии с так называемым четырехканальным сигнальным форматом *B*, который представляется четырьмя компонентами (сферическими функциями):

$$W = S/\sqrt{2},$$

$$X = S \cos \varphi \cos \theta,$$

$$Y = S \sin \varphi \cos \theta,$$

$$Z = S \sin \theta,$$

где *W* – сигнал ненаправленного микрофона; *X*, *Y*, *Z* – сигналы направленных микрофонов с диаграммами в виде «восьмерки», направленными вдоль координатных осей *x*, *y*, *z*, соответственно; φ – угол в горизонтальной плоскости *x*, *y*; а θ – угол возвышения относительно оси *z*.

Вид сферических функций, соответствующих диаграммам направленности, приведён на рисунке 18, где функция нулевого порядка соответствует компоненте *W*, а функции первого порядка компонентам *X*, *Y*, *Z*.

Описанный, четырёхканальный формат – это всего лишь простая форма *B*-формата первого порядка. *Ambisonics* второго порядка учитывает все функции, показанные на рисунке 18, т. е. использует 9 каналов передачи информации о направлении источников звука. *Ambisonics* третьего порядка увеличивает число каналов до 16, а четвёртого – до 49. Чем выше порядок *B*-формата, тем выше точность локализации источников звука.

Но даже формат первого порядка локализует звук лучше, чем традиционные технологии стереозвuka. В частности, с использованием *Ambisonics* отсутствует проблема «перескакивания» звука в системах виртуальной реальности с одного громкоговорителя (динамика) на другой, при повороте головы пользователя (вращении звукового поля). В системе *Ambisonics* предлагается система матричного кодирования для всех этапов обработки звука: *A*-формат – для записи сигналов микрофонами; *B*-формат для студийной обработки звука; *C*-формат для передачи звука и *D*-формат для воспроизведения.



Рисунок 31 –
Микрофон
типа *SoundField*

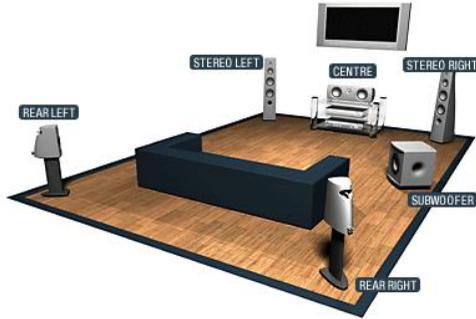


Рисунок 32 – Размещение динамиков для
surround (окружающего) звука формата 5.1:
3 фронтальных, 2 тыловых и сабвуфер

Подчеркнем, что *Ambisonics* можно декодировать на любой массив динамиков, в частности, на 2 динамика стереозвуча, 6 динамиков формата 5.1 (рисунок 32) или восемь динамиков 7.1 окружающего (в плоскости) звука; в том числе, на динамики, расположенные на сфере выше и ниже положения слушателя. С другой стороны записи в форматах: моно, стерео, 5.1 или 7.1 можно конвертировать в *Ambisonics*, например, с помощью плагина *Waves 360 Ambisonics*. На рисунке 33 показан интерфейс этого плагина. С помощью его элементов управления можно позиционировать (панорамировать) каждый элемент в звуковом поле.



Рисунок 33 – Кодер *Waves 360 Ambisonics*

Обычно звук воспринимается пользователем виртуальной реальности через наушники наголовного дисплея. Пространственный звук стал возможен благодаря технологиям *бинаурального звука*. Бинауральный процессор получает аудиовход и направление, в котором его следует расположить. Процессор добавляет к исходному сигналу слуховые сигналы так, что при воспроизведении в наушниках, пользователь воспринимает их в установленной виртуальной позиции. В итоге он воспринимает иммерсивное (виртуальное) звуковое поле, воспроизводимое в наушниках, при любом положении его головы. Несмотря на поворот головы слушателя виртуальные источники звука остаются на своем месте в пространстве. Примером технологии бинаурального звука (рисунок 34) является звуковое сопровождение смартфона *360 Reality Audio*, с использованием объектно-ориентированной стереофонической технологии *Sony 360⁰*.



Рисунок 34 – Иллюстрация ощущения иммерсионного звукового поля *360 Reality Audio* с направлениями на источники звука в полной сфере

Информация о местоположении «прикрепляется» к каждому звуку, например, вокалу, музыкальным инструментам, звуковым эффектам и человеческим голосам и размещается в сферическом пространстве. С помощью приложения для наушников на смартфоне фотографируется форма уха и анализируются слуховые характеристики пользователя, чтобы оптимизировать для него своё звуковое поле. Определение направления на источник звука представляет собой сложную задачу, которую мозг слушателя

учится решать в процессе развития не только с учётом разности фаз и громкости звука, приходящего на разные уши, но и с учётом формы его головы и ушных раковин и их положений в пространстве. Например, форма ушных раковин по-разному меняет спектр приходящего на улитку внутреннего уха звука у различных людей. Проводилось множество опытов в рамках изучения бинаурального слышания, при которых человеку вставляли небольшие микрофоны в слуховые проходы для записи звуков, приходящих с различных направлений. Далее такую стерео-запись воспроизводили этому же человеку через стереонаушники и человек различал направления на звуки. Однако другой человек с другой формой ушей не мог во многих случаях правильно определить направление записанного звука. Выяснилось, что попытки записывать звук с помощью микрофонов в ушах головы человека или специального манекена – *dummy head* (рисунок 35) и воспроизведение его в наушниках разных людей часто не дают нужного результата с точки зрения фиксации звуковых направлений.

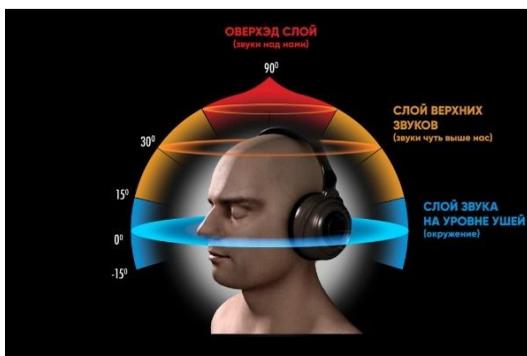


Рисунок 35 – Бинауральная запись звука; слева голова-манекен (*dummy head*), справа условное деление на «слои» звука по высоте положения источника

Для эффективного звуковоспроизведения в наушниках следует учитывать также форму плеч, шеи и тела слушателя. Звуковое поле формируется в условиях многочисленных ревербераций (отражений) от рельефа местности записи и на картину воспроизведения этого поля не должны накладываться реверберации помещения воспроизведения звука. В этом плане

использование наушников при воспроизведении представляются хорошим вариантом.

Возвращаясь к технологии *Ambisonics*, заметим, что она не даёт информации о дистанции до источника звука, а только о направлении на него из точки приёма. По-видимому, дистанцию до источника звука можно определить триангуляционным методом, сравнивая записи двух разнесённых микрофонов типа *Soundfield*, каждый из которых даёт информацию о направлении на источник. Здесь мы сталкиваемся с известной задачей пассивной пеленгации источника излучений. Стоит сказать, что мозг человека, формируя единую картину мира (в том числе виртуального) использует сигналы всех органов чувств человека. В частности, если человек одновременно видит и слышит источник звука, ориентируясь на зрительные данные о дистанции до источника, он может воспринимать звук, как приходящий с видимой дистанции, игнорируя несовпадение информации о дальности, полученной по звуковому и зрительному каналам.

В принципе, возможно воссоздание первичного акустического поля путём синтеза его копии с помощью большого количества динамиков, размещённых, например, регулярно в одной плоскости. При этом каждый динамик воспроизводит сферическую элементарную волну с выбранной задержкой (фазой) воспроизведения. Опираясь на фазы динамиков, можно синтезировать в точке слушания суммарное акустическое поле заданной конфигурации. В общем случае, согласно интегралу Кирхгофа–Гельмгольца, звуковое давление полностью определяется *в объёме*, свободном от источников звука, если звуковое давление и скорость движения среды определены *во всех точках поверхности объёма*. Добавим, что нет необходимости использовать «все точки». Достаточно чтобы выбранные точки расположения динамиков отстояли друг от друга на расстоянии не большем, чем половина длины волны наивысшей гармоники звукового поля в соответствии с теоремой Котельникова. То есть речь идет о возможности синтеза волнового поля (*wave field synthesis*) ограниченного спектра. Например, при частоте дискретизации (для компакт дисков) 44,1 кГц половина длины волны звука в воздухе равна 0,76 см. Если бы мы воспроизводили звуковое поле внутри сферы с радиусом один метр, то нам

понадобилось бы 125000 динамиков диаметром 0,75 см. Понятно, что создать такую сферическую фазированную решетку динамиков не просто. Создаются опытные образцы линейных решеток с целью продемонстрировать линейный (без высоты) звук с произвольным заданием программным путём дистанции источников звука. На рисунке 36 приведена фотография такой решетки с ограниченным спектром звукового диапазона. Вероятно, синтез волнового поля не всегда оптимален по экономическим причинам.



Рисунок 36 – Массив громкоговорителей с волновым синтезом в Эйндховене (более 200 динамиков)

Поэтому, для поиска новых эффективных методов создания ощущений объёмного звука, нам следует обратить особое внимание на механизм ощущения звука человеком. В частности, на особенность работы механизма его улитки внутреннего уха. Позднее мы получим математическую модель работы базилярной мембраны улитки с демонстрацией так называемой волны Бекеши, которая показывает отличие работы нашего уха при восприятии звука от технического анализатора спектра колебаний.

2. Телевизионный лазерный локаатор

Перейдем от общего описания принципов создания объёмного телевидения к описанию способов работы основных составных частей канала объёмного телевидения. К числу таких частей относится объёмная телекамера и объёмный монитор. Элементы линии связи – третьей составной части канала – могут строиться аналогично элементам линий связи обычного (плоского) телевидения, которые постоянно совершенствуются по мере развития различных систем связи: наземной, спутниковой, кабельной и других возможных.

Объёмная телекамера (см. рисунок 8) должна формировать четырёхкомпонентный видеосигнал, содержащий информацию о цвете и удалённости каждого элемента наблюдаемой телекамерой сцены (объекта). Три компонента видеосигнала кодируют цвет, а четвертая глубину – дальность элемента.

В каждый момент времени объект наблюдения фиксируется под некоторым углом (ракурсом) наблюдения. В зависимости от положения наблюдателя и поворота объекта угол наблюдения может изменяться. Серия снимков объекта под различными углами в ряде случаев позволяет определить его пространственную форму. Такой метод определения характерен для фотограмметрии [24]. Минимально необходимое количество снимков объекта различных ракурсов равно двум, что соответствует стереоскопической съёмке с получением стереопары изображений объекта. Рассматривание стереопары с помощью стереоскопа позволяет человеку во многих случаях ощутить трёхмерный характер формы объекта наблюдения и сделать правильное заключение о форме его поверхности. Особенностью зрения человека является не только наличие механизма нейронального анализа величины диспаратности (различий) стереопары, образованной на сетчатках двух глаз, но и наличие механизма воображения, работающего на основе опыта (памяти) восприятия сходных с наблюдаемым объектом. Последнее свойство зрения (мозга человека) не используется в автоматической фотограмметрии [25], нацеленной на быструю фиксацию новых пространственных форм. При автоматическом восстановлении (моделировании) формы объекта используются алгоритмы выявления соответствующих точек на обоих изображениях стереопары. При решении задач идентификации объектов с известными вариантами формы появляется возможность распознавать такие

объекты с помощью оптимальных пространственных фильтров, в частности с использованием технологий обучаемых нейронных сетей. Подчеркнем, что при любом способе многоракурсного воспроизведения формы объекта, необходимо наличие достаточного количества отождествляемых (соответствующих) точек на изображениях анализируемой серии фотоснимков и предположения о возможной форме объекта. В частности, этим способом не представляется возможным определить форму однородно окрашенного и освещённого объекта наблюдения, как это известно из стереоскопии [26]. Объекты наблюдения часто находятся на большом удалении от точки наблюдения и для получения качественной стереопары необходима достаточно большая база стереоскопического наблюдения. При относительном движении объекта и наблюдателя такая база может быть получена с задержкой на время их относительного смещения или поворота (оглядывания) объекта на необходимую величину. Наличие временной задержки при определении формы объектов в ряде случаев может представляться нежелательным фактором.

Отмеченные недостатки многоракурсных методов определения пространственной формы объектов наблюдения могут быть преодолены с использованием способа определения пространственной формы объектов [10], который назовем способом телевизионной лазерной локации по названию устройства, первой его реализации – телевизионного лазерного локатора (ТВЛЛ).

2.1. Способ телевизионной лазерной локации

Способ телевизионной лазерной локации, разработанный во Всесоюзном научно-исследовательском институте телевидения (Россия, 1980 г.), позволяет получить карту глубины удалённого объекта наблюдения за двойное время пролёта одного импульса света от точки наблюдения до объекта. Разрешение по углу изображения карты глубины при этом ограничено лишь разрешением телевизионного (фото) датчика, а разрешение по глубине может составлять единицы сантиметров.

На рисунке 37 приведена функциональная схема телевизионного лазерного локатора. Источник света 1 широким пучком подсвечивает поверхность трёхмерного объекта 2 короткими импульсами излучения. Отражённое от поверхности излучение поступает через оптическую систему 3, состоящую из объектива и разделяющей поток излучения пары зеркал, на два фотодатчика

(приёмника излучения). Датчик 4, например, ПЗС или КМОП матрица, формирует на своем выходе обычное изображение объекта пропорционально экспозиции излучения, прошедшего через полупрозрачное зеркало на каждый из её элементов. Другой датчик 5, например, электронно-оптический преобразователь (ЭОП), используя излучение, отражённое от пары зеркал, также формирует на своем выходе (экране) изображение.

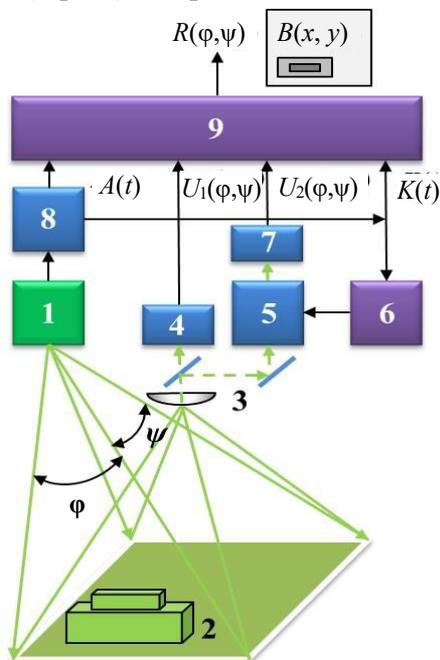


Рисунок 37 – Функциональная схема телевизионного лазерного локатора

Сформированное на экране ЭОП изображение – необычное, поскольку во время прихода отражённого от объекта излучения ЭОП быстро меняет (модулирует) своё усиление при воздействии сигнала блока управления 6. Например, при линейном увеличении усиления ЭОП в процессе приёма излучения дальние элементы поверхности объекта на экране ЭОП будут выглядеть более яркими по сравнению с ближними при одинаковом коэффициенте отражения поверхности, поскольку импульс света от них приходит позже, когда усиление ЭОП уже увеличилось. Необычность изображения на экране ЭОП обусловлена одновременным воздействием двух факторов на его формирование: тона раскраски и рельефа поверхности объекта наблюдения. Яркость элемента

изображения определяется как коэффициентом отражения импульса света от элемента поверхности объекта, так и запаздыванием прихода этого импульса на приёмник пропорционально удалённости элемента поверхности. Подчеркнем, что запаздывание влияет на яркость элемента изображения в основном из-за модуляции усиления ЭОП. Изображение с экрана ЭОП репродуцируется сочлененным с ним фотодатчиком 7, например, КМОП матрицей, обеспечивающей его последующую развёртку. Момент излучения импульса света источника подсветки фиксируется специальным быстродействующим одноэлементным фотодатчиком 8, сигнал которого запускает счётчик времени задержки начала формирования сигнала модуляции ЭОП. Система формирования карты глубины рельефа поверхности объекта наблюдения 9 обеспечивает поэлементное вычисление отношения видеосигналов соответствующих элементов (сенселей) фотодатчиков 7 и 4. Полученная матрица отношений $R(x, y)$ с учётом формы импульсов излучения подсветки и закона модуляции усиления ЭОП служит основой для формирования карты глубины $B(x, y)$ удалённого объекта наблюдения.

Для синтеза цветного объёмного изображения объекта наблюдения необходима информация о цвете и удалённости каждого элемента его видимой поверхности в точке наблюдения. Использование ТВЛЛ в сочетании с цветной телевизионной телекамерой обеспечивает получение такой информации. При этом исходная информация о цвете элементов объекта кодируется как обычно тремя R (красная), G (зелёная) и B (синяя) компонентами полного изображения, а информация о дальности этих элементов дополнительной компонентой дальности D . Сегодня эту компоненту в литературе по обработке изображений принято называть картой глубины (*depth map*). Полное $RGBD$ изображение из четырёх компонент, приведённое на рисунке 8, может быть воспроизведено на четырёх одноцветных мониторах или на одном цветном стереоскопическом после синтеза из него стереопары для его визуального анализа. Для быстрой идентификации объектов наблюдения следует использовать автоматическое распознавание их образов непосредственно с использованием R , G , B и D компонент. Объёмную телекамеру с использованием ТВЛЛ принято называть *временнóлетной* (англ. *Time of Flight*, ToF) в соответствии с классификацией, принятой в [11]. Этот термин соответствует определению дальности до точек поверхности объектов наблюдения по времени запаздывания прихода к телекамере, отражённого от них импульса подсветки.

Способ ТВЛЛ может использоваться для получения телевизионного изображения карт глубины с частотой кадров, например, 50 Гц или отдельных карт глубины объектов без создания видеопотока. В качестве импульсных источников света для освещения объектов наблюдения удобно использовать лазеры подходящего типа.

2.2. Математическая модель ТВЛЛ

Приведём математическое описание способа телевизионной лазерной локации. Сигнал датчика 4 (см. рисунок 37) можно представить выражением:

$$U_1(\varphi, \theta) = M \frac{I(\varphi, \theta) \rho(\varphi, \theta)}{f[r^2(\varphi, \theta)]} \int_{t_1}^{t_2} A \left[t - \frac{2r(\varphi, \theta)}{c} \right] dt, \quad (1)$$

где φ – горизонтальная, а θ – вертикальная угловые координаты точки поверхности трёхмерного объекта (начало координат, излучатель и фотодатчик находятся в одной точке пространства); M – постоянная; $I(\varphi, \theta)$ – угловое распределение интенсивности лазерного излучения; $\rho(\varphi, \theta)$ – распределение коэффициента отражения поверхности объекта; $r(\varphi, \theta)$ – координата дальности точки поверхности объекта, $f[r^2(\varphi, \theta)]$ – функция учёта зависимости освещённости объекта от расстояния до него; $A(t)$ – форма импульса лазерного излучения; t – время; c – скорость распространения лазерного излучения. Пределы интегрирования t_1 и t_2 взяты с учётом конечности длительности импульса лазерного излучения и расстояния до объекта наблюдения.

При условии нормировки (соответствующем выборе значения постоянной M):

$$\int_{t_1}^{t_2} A \left[t - \frac{2r(\varphi, \theta)}{c} \right] dt = 1. \quad (2)$$

При этом:
$$U_1(\varphi, \theta) = M \frac{I(\varphi, \theta) \rho(\varphi, \theta)}{f[r^2(\varphi, \theta)]}. \quad (3)$$

Сигнал датчика 7 (см. рисунок 37) можно представить как:

$$U_2(\varphi, \theta) = M \frac{I(\varphi, \theta) \rho(\varphi, \theta)}{f[r^2(\varphi, \theta)]} \int_{t_1}^{t_2} A \left[t - \frac{2r(\varphi, \theta)}{c} \right] K(t) dt, \quad (4)$$

где $K(t)$ – выбранный закон модуляции ЭОП.

С учётом выражений (1) и (4) отношение сигналов фотодатчиков 7 и 4 определяется равенством:

$$O(\varphi, \theta) = \frac{U_2(\varphi, \theta)}{U_1(\varphi, \theta)} = \int_{t_1}^{t_2} A \left[t - \frac{2r(\varphi, \theta)}{c} \right] K(t) dt. \quad (5)$$

Если $K(t)$ и $A(t)$ выбраны так, что правая часть этого равенства взаимно-однозначная функция от $r(\varphi, \theta)$ на выбранном интервале дальности, то и $O(\varphi, \theta) = O[r(\varphi, \theta)] = O(r)$ – взаимно-однозначная функция дальности на данном интервале дальностей. При этом можно однозначно определить пространственную форму удалённого объекта:

$$R(\varphi, \theta) = r(\varphi, \theta) - r_0, \quad (6)$$

где r_0 – дальность до ближайшей точки объекта.

Определить значение r_0 можно различными способами, сводящимися к определению минимума функции $r(\varphi, \theta)$. Приведённое ограничение на выбор функций $K(t)$ и $A(t)$ допускает большое разнообразие вариантов их возможного выбора. Рассмотрим некоторые из этих вариантов.

2.3. Однозональный вариант ТВЛЛ

При предельно малой длительности, но конечной энергии импульса излучения лазера его можно представить функцией Дирака $A(t) = \delta(t)$. Выберем линейный закон модуляции усиления ЭОП $K(t) = at$. Тогда, подставляя эти значения в выражение (5) и используя фильтрующее свойство δ -функции, получим:

$$O(\varphi, \theta) = a \int_{t_1}^{t_2} \delta \left[t - \frac{2r(\varphi, \theta)}{c} \right] t dt = \frac{2ar(\varphi, \theta)}{c}, \quad (7)$$

где a – постоянная.

При этом пространственная форма объекта – карта глубины – определяется простым выражением:

$$R(\varphi, \theta) = \frac{c}{2a} O(\varphi, \theta) - r_0. \quad (8)$$

На рисунке 38 приведены прямая и инверсная карты глубины пространственной формы сферического объекта наблюдения; под каждой картой глубины приведены соответствующие формы светового импульса $A(t)$, сигнала модуляции $K(t)$ и сигнала отношения $O(r)$ с приведением графиков к выбранному значению углов φ, θ .

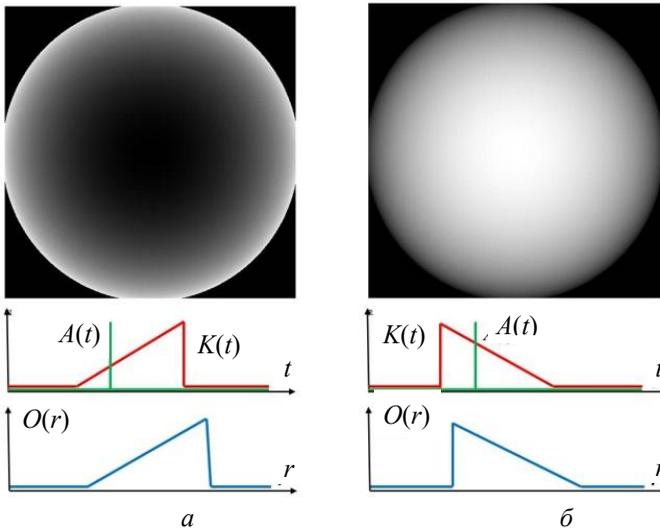


Рисунок 38 – Карты глубины пространственной формы сферического объекта наблюдения и соответствующие им графики функций $A(t)$, $K(t)$ и $O(r)$: прямая – а; инверсная – б

Карта глубины представляет собой изображение, яркость которого:

$$B(x, y) = \beta R(\varphi, \theta), \quad (9)$$

где $x = \eta\varphi$, $y = \eta\theta$; β и η – постоянные.

Инверсную карту глубины можно определить выражением:

$$B_1(x, y) = B_{\max}(x, y) - B(x, y), \quad (10)$$

где $B_{\max}(x, y)$ – максимальное значение $B_{\max}(x, y)$.

Важной характеристикой ТВЛЛ является его разрешающая способность. Угловая разрешающая способность ТВЛЛ определяется разрешающей способностью его телевизионных фотодатчиков. Разрешающая способность ТВЛЛ по глубине в основном определяется уровнем шумов этих фотодатчиков, которые, в свою очередь, определяются значением дисперсии их фотонных или дробовых (для ЭОП) шумов. Приведём оценку разрешающей способности ТВЛЛ по глубине по уровню фотонных шумов. Максимальное число уровней квантования сигнала фотодатчика (элементов разрешения по глубине) с неравномерным шагом квантования, равном двукратному значению дисперсии фотонного шума:

$$k_{\max} = \sqrt{n_{\max} + 7} - 3, \quad (11)$$

где n_{\max} – максимальное число фотонов на сенсель фотодатчика.

Например, для 1000 фотонов, приходящихся на сенсель, $k_{\max} \approx 29$, для 10000 фотонов – $k_{\max} \approx 97$.

Число элементов разрешения по глубине, согласно выражению (11) приходится на всю зону по глубине, определяемую длительностью сигнала модуляции ТВЛЛ, как приведённый на графиках рисунка 38. Последнее означает, что для увеличения числа элементов разрешения по глубине поверхности объекта наблюдения следует максимально сужать размер зоны наблюдения по глубине, уменьшая длительность сигнала модуляции, но сохраняя его размах для максимально возможного отношения видеосигналов фотодатчиков $O(r)$.

2.4. Многозональный вариант ТВЛЛ

Для увеличения разрешения ТВЛЛ по глубине можно наряду с уменьшением размера зоны наблюдения по глубине увеличить число зон, приходящихся на рельеф поверхности объекта.

На рисунке 39 приведены прямая и инверсная многозональные карты глубины объекта наблюдения со сферической формой поверхности и соответствующие графики сигналов модуляции и отношения видеосигналов фотодатчиков $O(r)$.

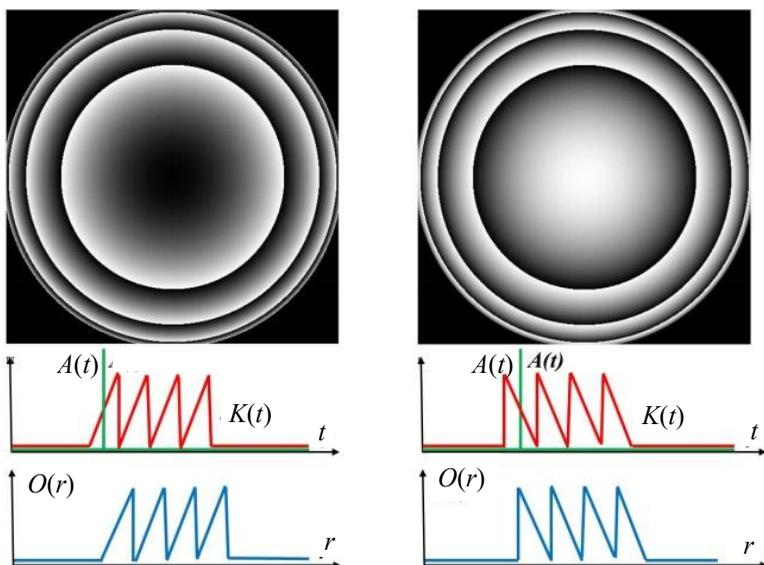


Рисунок 39 – Многозональные карты глубины пространственной формы сферического объекта наблюдения и графики функций $A(t)$, $K(t)$ и $O(r)$:
прямая – а; инверсная – б

Увеличение числа элементов разрешения ТВЛЛ по глубине прямо пропорционально числу зон разрешения, приходящихся на глубину рельефа поверхности объекта. Однако, при многозональном режиме работы в общем случае возникает проблема отождествления принадлежности данного элемента поверхности объекта соответствующей зоне наблюдения. Такое отождествление можно выполнять, используя межзональные границы на картах глубины, которые, выполняя роль изолиний на топографических картах, позволяют производить определение номера зоны наблюдения для данного элемента поверхности. После отождествления появляется возможность построения единой карты глубины поверхности объекта с высоким разрешением по глубине без разделения её на зоны глубины. Решение задачи отождествления зон и синтеза цифровой модели рельефа поверхности объекта наблюдения требует разработки соответствующего программного обеспечения. Рассмотренные выше варианты ТВЛЛ предполагают использование коротких мощных импульсов лазерного излучения.

2.5. Импульсный лазер

На рисунке 40 приведена функциональная схема импульсного твердотельного лазера с полупроводниковой торцевой накачкой.

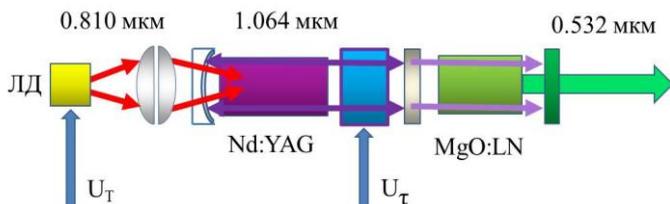


Рисунок 40 – Функциональная схема импульсного твердотельного лазера с полупроводниковой торцевой накачкой

Эти лазеры, обладая небольшими габаритами и хорошим ресурсом, способны генерировать мощные (единицы мегаватт) короткие (единицы наносекунд) импульсы света с частотой телевизионных кадров. Такие характеристики являются идеальными для освещения удалённых объектов, движущихся с любыми реальными скоростями. Согласно функциональной схеме во время действия токового импульса U_T , импульс излучения лазерного диода ЛД через концентрирующие линзы и прозрачное для длины волны $0,810 \text{ мкм}$ зеркало резонатора лазера попадает через торец в активный элемент гранатового лазера, осуществляя его накачку.

По заднему фронту электрического импульса U_T оптического затвора происходит лавинообразное высвечивание накопленной энергии активного элемента в виде короткого мощного импульса инфракрасного (ИК) излучения с длиной волны 1,064 мкм. Этот импульс через полупрозрачное выходное зеркало резонатора лазера попадает в нелинейный оптический элемент, где значительная часть его энергии преобразуется в импульс излучения с длиной волны 0,532 мкм (зелёного цвета), который проходит через зелёный светофильтр, подавляющий ИК излучение. Преобразование длины волны света лазера может быть полезным для получения карт глубины, при работе, например, в водной среде, где ИК излучение быстро затухает, кроме того, некоторые современные фотокатоды ЭОП значительно более эффективны в зеленой области спектра, чем в ИК диапазоне [27]. Вместе с тем во многих случаях полезна работа с невидимым глазу ИК излучением для получения карт глубины. При этом нет необходимости в присутствии в составе лазера секции удвоения частоты излучения (уменьшения длины волны до 0,532 мкм).

Импульс излучения лазера можно моделировать зависимостью:

$$I(t) = \frac{b^2 t e^{-bt/T}}{T^2}, \quad (12)$$

где $b = 1,8$; $T = (2 \dots 10)$ нс – длительность импульса на уровне 0,7; $t = (0 \dots \infty)$. В интервале времени импульса от 0 до T сосредоточена половина его энергии.

2.6. Приём отраженного излучения лазера

В светлое время суток телевизионный лазерный локатор будет вынужден работать в условиях интенсивной солнечной засветки. Средняя энергия отражённого от объекта наблюдения излучения Солнца при этом на много порядков превосходит среднюю энергию отражённого лазерного излучения. Для обеспечения нормальной работы ТВЛЛ следует организовать его действие таким образом, чтобы экспозиция отражённого от объекта солнечного излучения на его фотодатчиках была во много раз меньше, чем экспозиция отражённого лазерного излучения. Поскольку экспозиция пропорциональна времени действия излучения, следует оснастить фотодатчики электронными затворами, открываемыми лишь на время прихода на фотодатчики коротких импульсов отражённого лазерного излучения. При использовании такого режима работы

ТВЛЛ солнечное излучение будет восприниматься лишь в течении времени, не превышающего, например, 20 наносекунд. В этом случае при пятидесяти лазерных импульсах в секунду, мощность солнечного излучения будет подавляться в миллион раз. Возможность достижения подобных коэффициентов подавления является наиболее привлекательной характеристикой одноимпульсного (на телевизионный кадр) варианта ТВЛЛ. На рисунке 41 приведена функциональная схема ЭОП с микроканальной пластиной (МКП) и зависимость коэффициента усиления ЭОП от напряжения на МКП.

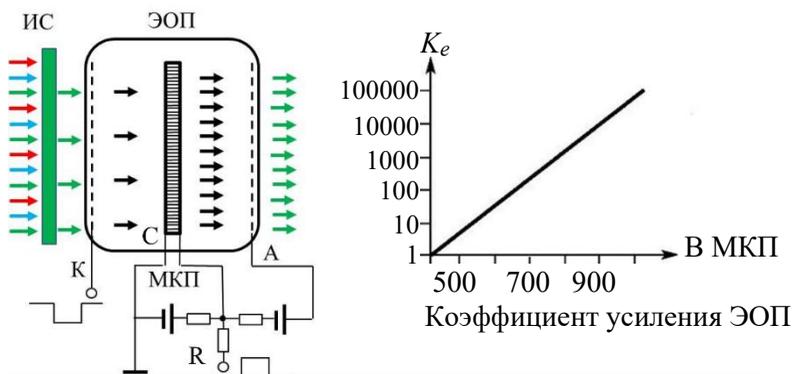


Рисунок 41 – Функциональная схема электронно-оптического преобразователя и зависимость его коэффициента усиления от напряжения на микроканальной пластине

При подаче на МКП прямоугольного импульса напряжения через времязадающий резистор R , происходит заряд ёмкости C МКП

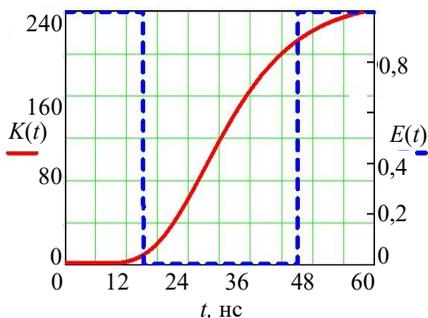


Рисунок 42 – Зависимость коэффициента усиления ЭОП от времени и график работы затвора

по экспоненциальному закону с постоянной времени $\tau = RC$. Коэффициент усиления ЭОП, согласно графику на рисунке 41, зависит от напряжения на МКП по логарифмическому закону. Итоговая зависимость коэффициента усиления ЭОП от времени $K(t)$ показана на рисунке 42. Здесь же приведён график работы затвора ЭОП $E(t)$ в условных единицах. Можно видеть, что во время открытого затвора

зависимость коэффициента усиления ЭОП от времени носит практически линейный характер. Для организации режима затвора ЭОП удобно использовать подачу импульса напряжения на промежуток катод K -МКП, как показано на рисунке 41, поскольку для этого достаточно импульса сравнительно небольшого напряжения в несколько сотен вольт. Поскольку ЭОП является чувствительным к свету прибором, работающем в узком диапазоне длин волн лазерного излучения, можно ослабить действие на него солнечного излучения, установив перед его фотокатодом интерференционный светофильтр – ИС (см. рисунок 41) с полосой пропускания, соответствующей длине волны лазерного излучения.

2.7. Многоимпульсные варианты ТВЛЛ

Выше был описан одноимпульсный режим работы ТВЛЛ, когда для получения одного телевизионный кадра используется один импульс подсветки объекта. В ряде приложений ТВЛЛ может оказаться удобным использование его многоимпульсного варианта, при котором объект подсвечивается множеством сравнительно слабых импульсов излучения. Энергия отражённых импульсов суммируется фотоприёмниками ТВЛЛ в течении времени кадра. Обязательным условием в этом случае является синхронность принимаемых импульсов излучения и импульсов модуляции усиления фотоприёмника. В условиях отсутствия существенных паразитных засветок дешевле использовать лазер непрерывного излучения с внешним модулятором. Например, при создании действующего макета ТВЛЛ в НИИ телевидения использовалось синусоидально модулированное лазерное излучение и близкая к синусоидальной модуляция МКП ЭОП. При этом модулятор лазера и МКП питались от одного генератора с установленной частотой 100 МГц. На рисунке 43 приведена фотография этого макета. В правом нижнем углу рисунка приведены карты глубины зондируемого тест-объекта в прямом и инверсном виде, которые воспроизводились на экране чёрно-белого монитора. При указанной частоте модуляции пространство объекта разбивается на зоны глубины величиной по семьдесят пять сантиметров вдоль оптической оси ТВЛЛ. При переводе тест-объекта, установленного на специальные рельсы, из одной зоны в соседнюю его тональная раскраска на экране монитора менялась на инверсную. При синусоидальной модуляции между зонами образуются межзональные промежутки низкого разрешения по глубине и для их

уменьшения вместо синусоидальной желательнее использовать прямоугольную модуляцию света и коэффициента усиления фотоприемника ТВЛЛ в виде меандра.

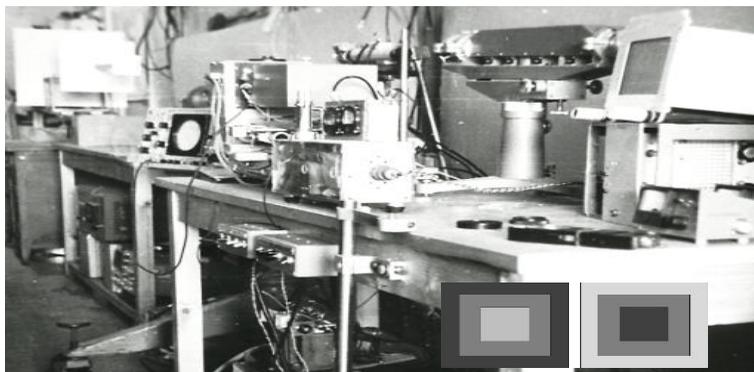


Рисунок 43 – Макет телевизионного лазерного локатора (ТВЛЛ) НИИ телевидения

На рисунке 44 приведены прямая и инверсная карты глубины сферического объекта при работе ТВЛЛ в режиме прямоугольной модуляции. В данном случае поверхность объекта попадает в четыре зоны глубины – две прямые и две инверсные.

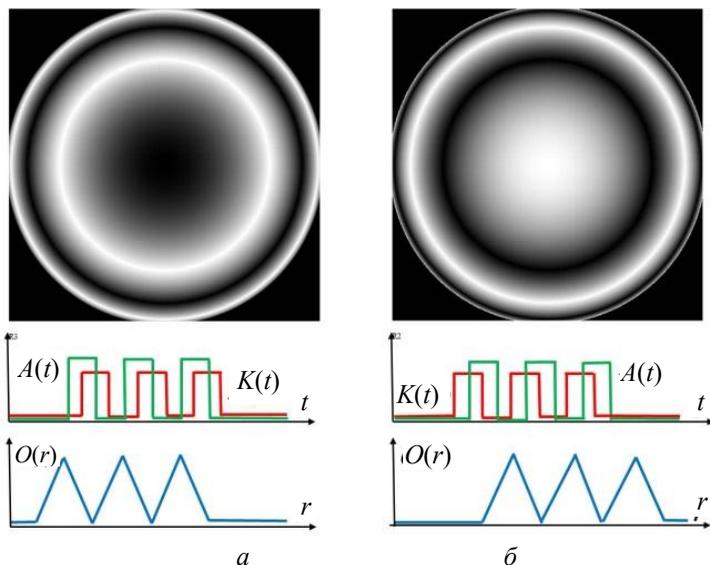


Рисунок 44 – Многозональные карты глубины сферического объекта и графики функций $A(t)$, $K(t)$ и $O(r)$ при многоимпульсном режиме ТВЛЛ и прямоугольной модуляции: прямая – a ; инверсная – $б$

Описанные многоимпульсные режимы работы ТВЛЛ принципиально отличаются от моноимпульсных тем, что в первом случае энергия света суммируется в фотоприёмнике, а во втором – в рабочем теле импульсного лазера. В любом случае этой энергии должно быть достаточно для формирования хорошего изображения в фотоприёмниках ТВЛЛ.

2.8. Пространственное суммирование подсветки

Вместо одного лазера с внешним модулятором для многоимпульсной подсветки объекта можно использовать решетки лазерных светодиодов, работающие в синхронном импульсном режиме, которые обеспечивают необходимую освещённость объекта. Такое решение использовалось в японском макете ТВЛЛ [28]. На рисунке 45 приведены фотографии макета и его разработчика Масахиро Кавакиты, а также карта глубины автора разработки, полученная с использованием макета.

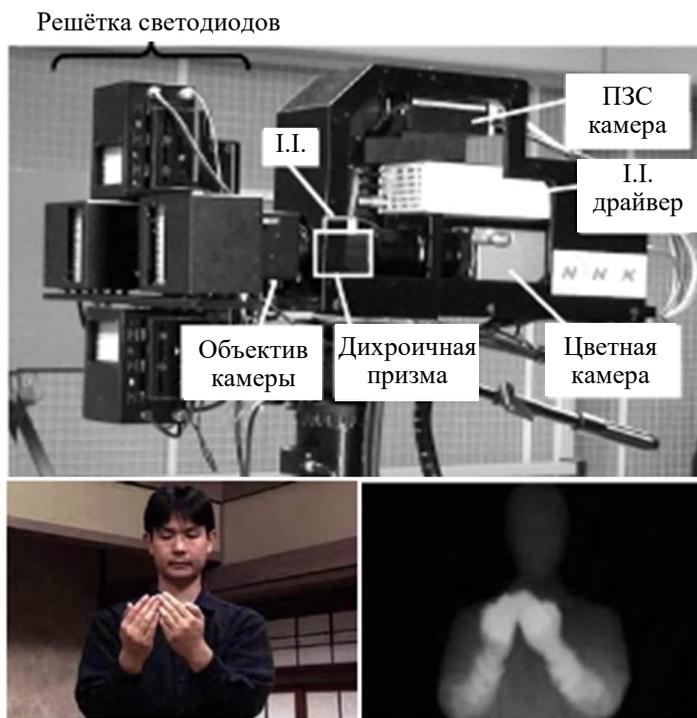


Рисунок 45 – Японский макет ТВЛЛ, сочленённый с цветной телекамерой высокой четкости (вверху), изображение разработчика и его карта глубины (внизу)

В макете применялась равнобедренная пирамидальная модуляция излучения света с частотой от 10 до 50 МГц и прямоугольная модуляция усиления (затворный режим) фотодатчика – ЭОП с МКП с длительностью импульсов от 1 до 20 нс. Цветная телекамера высокой чёткости обеспечивала формирование цветных компонент изображения объекта. Таким образом макет обеспечивал формирование *RGBD* изображения объекта наблюдения, аналогичное показанному на рисунке 8. При использовании решеток светодиодов осуществлялось пространственное суммирование энергии подсветки объекта.

Теперь, после описания некоторых схем ТВЛЛ, приведём формулу изобретения [10], описывающую принцип их конструирования: *Способ определения пространственной формы объектов, включающий импульсный подсвет объекта, разделение отражённого излучения на два пучка и регистрацию изображения объекта в каждом пучке, отличающийся тем, что, с целью повышения разрешающей способности по дальности, модулируют интенсивность излучения одного из пучков в процессе регистрации изображения объекта и определяют форму объекта по известному закону модуляции и полученным распределениям освещенности в обоих изображениях.* Форма объекта, в частности, фиксируется картой глубины – компонентой дальности.

Отметим принципиальную возможность организации работы матричных фотодатчиков в режиме электронного затвора, что позволяет отказаться от использования в ТВЛЛ высоковольтных элементов ЭОП. Переход на КМОП датчики [68] позволяет избавиться от специфических дробовых шумов МКП и повысить таким образом разрешающую способность ТВЛЛ по глубине [69], [70]. Следует подумать о возможности совмещения всех функций *RGBD*-телекамеры в рамках единой матрицы фотодатчика. Принцип субпиксельной организации таких фотодатчиков подсказывает принципиальную возможность компактного анализа различных параметров светового излучения. При этом было бы интересно получить не только *R*, *G*, *B* цветные компоненты, но вообще любые компоненты спектрального деления [29] отражённого излучения, а также *D* компоненту дальности или компоненты вектора Стокса поляризации излучения [30] на каждом небольшом участке (пикселе) матрицы фотодатчика. Знание о виде поляризации излучения, отражённого от объекта наблюдения, может быть необходимым при наблюдении объектов, имеющих зеркальную компоненту отражения.

Заканчивая описание подхода к созданию активной объёмной телекамеры, нам следует перейти к описанию подхода к созданию объёмного монитора – другой важной части канала объёмного телевидения, который обеспечивал бы действие зрителя в телевизионном мире, реализуя принципы *объёмности* и *трёхмерной интерактивности*. Важным требованием к такому монитору является обеспечение согласования конвергенции и аккомодации глаз пользователя в процессе наблюдения.

3. Объёмный монитор открытого типа

На рисунке 4 представлена схема объёмного телевидения с телевизором открытого типа. В этом случае у телезрителя возникает иллюзия присутствия внутри телевизионного изображения. Подобную иллюзию может создать и наголовный дисплей виртуальной реальности, показанный на рисунке 14 а. Недостатком такого дисплея является несоответствие аккомодации и конвергенции глаз пользователя. На рисунке 4б приведён вид подобного дисплея (без смартфона), выполненного из картона и предназначенного для игрушки «Интерактивный глобус».



а



б



в

Рисунок 4б – Картонная версия наголовного объёмного дисплея (без смартфона): вид спереди – а; вид сзади со стороны лица пользователя – б; стереопара на экране смартфона – в

Несмотря на простоту конструкции, он позволяет, благодаря использованию в его составе смартфона, обеспечивающего формирование стереопары, получить стереоскопическое (объемное) изображение. Специальное программное обеспечение, устанавливаемое на смартфон в качестве приложения, позволяет пользователю погрузиться в одну из нескольких версий виртуальной среды. При этом поворачивая голову он может осуществлять круговой обзор, а совмещая поворотом головы специальный маркер, находящийся в поле зрения, со встроенными виртуальными органами управления, может перемещаться в среде или выполнять в ней возможные действия. Описанный вариант не требует наличия в качестве органа управления положением пользователя в виртуальной среде специального джойстика. Используется только встроенный в смартфон инерциальный датчик положения головы. Разумеется, показанный «игрушечный» стереоскоп не обладает в полном объеме функциями и качествами современного наголовного дисплея, однако обеспечивает приемлемое разрешение и динамику стереоскопического изображения. Но, как и у его дорогостоящего собрата, стереопара, наблюдаемая через линзы стереоскопа (см. рисунок 46 в), создает иллюзию наблюдения изображения на заданном расстоянии оптимальной аккомодации. При попытке пользователя изменить глубину наблюдения конвергенция его глаз изменяется на нужную дистанцию, а попытка соответствующего изменения аккомодации приводит к расфокусировке изображения. Такая рассогласованность естественной работы глаз при переносе точки внимания на различную глубину приводит к утомлению глаз.

3.1. Согласование конвергенции и аккомодации

Для согласования механизмов аккомодации и конвергенции глаз необходимо изменять радиус сферической волны света от наблюдаемого элемента объемного изображения в соответствии с его положением в виртуальном пространстве. Положение объекта в виртуальном пространстве определяется формулой лупы стереоскопа. На рисунке 47 приведена оптическая схема такой лупы, где R – расстояние до виртуального объекта (м); r – расстояние до экрана смартфона (м); f – фокусное расстояние лупы (м); γ – увеличение лупы; H и h – линейные размеры виртуального и физического изображения объекта на экране смартфона, соответственно

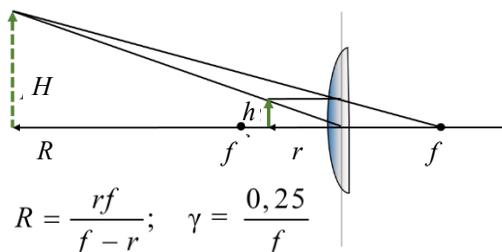


Рисунок 47 – Оптическая схема линзы стереоскопа

В соответствии с оптической схемой и свойствами подобных треугольников запишем равенства

$$\frac{r}{R} = \frac{h}{H}, \quad \frac{h}{H} = \frac{f}{f + R},$$

из сравнения которых получим выражение для расстояния от линзы до мнимого (виртуального) изображения стереоскопа

$$R = \frac{rf}{f - r}, \quad (13)$$

где f – фокусное расстояние линзы, r – расстояние от линзы до экрана смартфона.

На рисунке 47 также приведена формула для видимого увеличения линзы [31], определяемого как отношение угловых размеров объекта наблюдения в виртуальном пространстве и на экране смартфона с расстояния наилучшего зрения 0,25 м:

$$\gamma = \frac{0,25}{f}. \quad (14)$$

В соответствии с (13) видимое расстояние до виртуального объекта можно менять, изменяя либо расстояние от линзы до экрана, на котором воспроизводится плоское изображение, либо изменяя фокусное расстояние линзы. Первый вариант может быть реализован с использованием метода, использующего технологию *DepthCube*, показанного на рисунке 12 б, организуя проекцию плоского изображения на стопку экранов, которые прозрачны в обычном состоянии и один из которых активируется на нужном расстоянии от линзы, становясь рассеивающим свет проектора. В данном случае основной технологической проблемой является создание тонкой стопки экранов с достаточным их числом для организации дискретного управления положением активного экрана.

Второй вариант предполагает использование в качестве лупы стереоскопа оптической системы с изменяемым (управляемым) фокусным расстоянием и единственным экраном, в качестве которого можно использовать экран смартфона. На рисунке 48 приведена оптическая схема такой системы, которая включает две линзы: основную лупу стереоскопа с фокусным расстоянием f_1 и так называемую гидравлическую линзу с фокусным расстоянием f_2 .

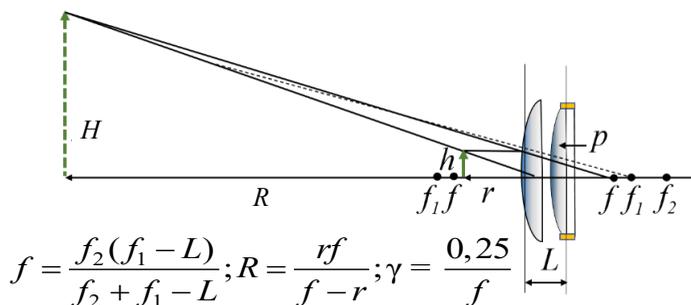


Рисунок 48 – Оптическая схема системы с управляемым фокусным расстоянием, где R – расстояние до виртуального объекта (м); r – расстояние до экрана смартфона (м); f – фокусное расстояние оптической системы из двух линз (м); f_1 – фокусное расстояние лупы стереоскопа; f_2 – фокусное расстояние линзы с управляемым фокусным расстоянием (гидролинзы); L – расстояние между линзами; p – давление гидролинзы; γ – увеличение системы; H и h – линейные размеры виртуального и физического изображения объекта на экране смартфона, соответственно

Гидравлическая линза представляет собой короткий цилиндр, заполненный прозрачной жидкостью с высоким коэффициентом преломления, например, глицерином. Основания цилиндра выполнены из прозрачного материала, в качестве которого в одном или обоих случаях используется упругая пленка. Внутренняя полость цилиндра соединена с миникомпрессором, задающим давление в цилиндре. При повышении давления упругая пленка цилиндра выгибается, приобретая сферическую поверхность, увеличивая оптическую силу гидролинзы. Оптическая система из двух собирающих линз имеет фокусное расстояние [31]

$$f = \frac{f_2(f_1 - L)}{f_2 + f_1 - L}, \quad (15)$$

где L – расстояние между линзами.

Подставляя выражение (15) в (13), полагая $f_2 \rightarrow \infty$ и задавая минимальное расстояние до объектов наблюдения R_{\min} , найдем необходимое положение экрана смартфона в стереоскопе

$$r = \frac{1}{\frac{1}{R_{\min}} + \frac{1}{f_1 - L}}. \quad (16)$$

Полагая в (15) $f_2 = 1/D_\Gamma$, где D_Γ – оптическая сила гидравлической линзы и подставляя в (13) значения из (15) и (16), получим общее выражение для расстояния от оптической системы до мнимого (виртуального) изображения стереоскопа, оснащённого гидравлической линзой

$$R(D_\Gamma) = \frac{1}{\frac{1}{R_{\min}} - D_\Gamma}, \quad (17)$$

где $R(D_\Gamma) > R_{\min}$.

На рисунке 49 приведён график зависимости (17), при $R_{\min} = 0,6$ метра, $f_1 = 0,046$ метра, $r = 0,04$ метра и $L = 0,003$ метра.

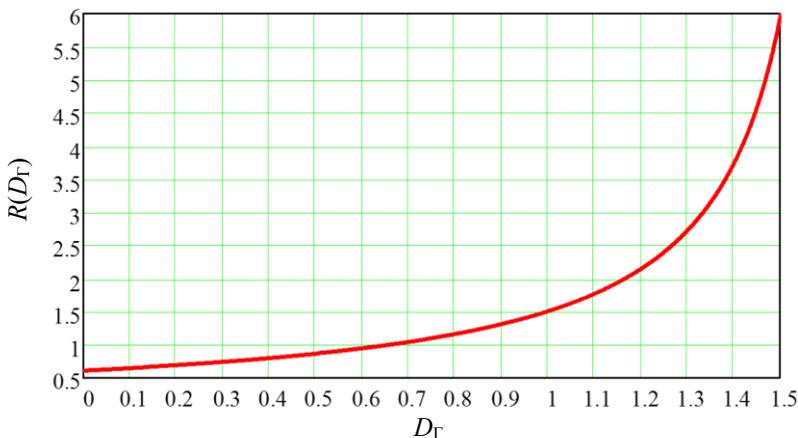


Рисунок 49 – Дистанция наблюдения изображения в стереоскопе в зависимости от оптической силы гидравлической линзы

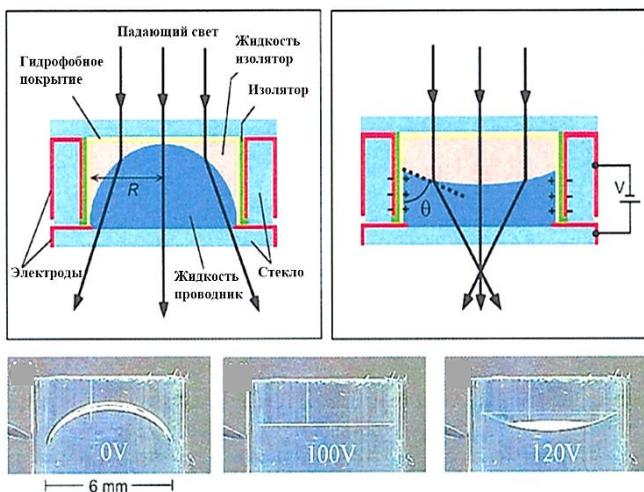
Оптическая сила гидравлической линзы D_Γ выражена в диоптриях (дптр), расстояние до объекта в метрах; по графику видно, что на дистанциях от 0,6 до 6 метров оптическую силу гидравлической линзы следует изменять в пределах от 0 до 1,5 диоптрии. Перестройку оптической силы в приведённом диапазоне расстояний

необходимо производить с достаточной скоростью в течении долей секунды для того, чтобы пользователь не ощущал запаздывания реакции дисплея на изменение дистанции наблюдения. Перестраиваемые линзы в настоящее время используются, например, в адаптивных очках *ViXion 01* японской компании *ViXion Inc* с автоматической регулировкой фокуса (рисунок 50) [32], где время реакции находится в пределах одной секунды. В данном случае автоматическая регулировка достигается использованием *ToF* (времяпролётной) технологии для определения расстояния до объекта наблюдения. Ёмкость аккумулятора 150 мАч обеспечивает 10 часов непрерывной работы очков, при времени зарядки 3 часа. Фокусное расстояние от 0,05 м (20 дптр) до бесконечности (0 дптр).

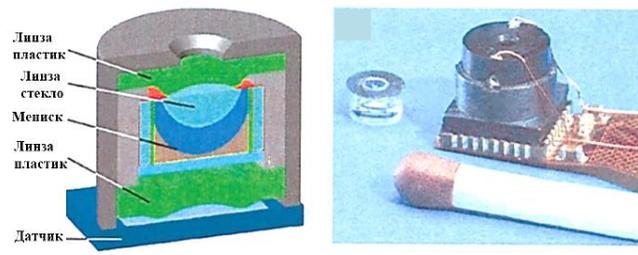


Рисунок 50 – Адаптивные очки *ViXion 01* с автоматической регулировкой фокуса

В общем случае скорость перестройки оптической силы линзы определяется как параметрами самой линзы, так и параметрами устройства управления. Способы перестройки линзы зависят от используемой для её создания технологии. Например, в качестве оптической линзы может использоваться мениск между двумя несмешивающимися жидкостями. Изменение кривизны этого мениска за счёт электросмачивания при изменении прилагаемого к нему напряжения приводит к изменению фокусного расстояния линзы [33]. На рисунке 51 приведены иллюстрации из [33], поясняющие принцип действия данного способа, который применялся для изменения фокусного расстояния миниатюрной телекамеры. Недостатки способа – в малом размере линзы и высоком напряжении управления.



a



б

Рисунок 51 – Пример оптической линзы, изменяющей кривизну мениска между двумя несмешивающимися жидкостями: принцип действия – *a*; конструкция и внешний вид электролинзы – *б*

Гидравлическая линза лишена этих недостатков. Такая линза может иметь мембрану как с однородной, так и неоднородной толщиной для уменьшения сферической aberrации [34]. Временные параметры гидравлической линзы зависят, в частности, от изменения объёма (массы) жидкости, с которой оперирует управляющий микрокомпрессор и его мощности.

Максимальное изменение объёма гидролинзы соответствует объёму шарового сегмента с радиусом основания r_0 и радиусом сферы гидролинзы $r_T = (n - 1)/D_T$, где n – коэффициент преломления используемой жидкости. Изменение массы сегмента зависит от изменения его объёма и плотности жидкости ρ и определяется выражением

$$m(D_{\Gamma}) = \frac{\pi\rho}{3} \left\{ 2 \left(\frac{n-1}{D_{\Gamma}} \right)^3 - \sqrt{\left(\frac{n-1}{D_{\Gamma}} \right)^2 - r^2} \left[2 \left(\frac{n-1}{D_{\Gamma}} \right)^2 + r^2 \right] \right\}. \quad (18)$$

При $\rho = 1,25 \times 10^6 \text{ г/м}^3$ (глицерин), $r_0 = 0,012 \text{ м}$, $n = 1,47$ и изменении оптической силы линзы в пределах от 0 до 1,5 диоптрий, получим зависимость, приведённую на рисунке 52. Можно видеть, что в указанном диапазоне изменение массы не превышает

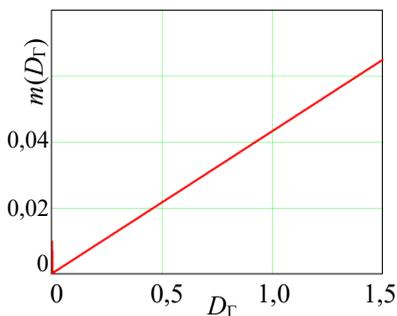


Рисунок 52 – Изменение массы жидкости гидролинзы в зависимости от её оптической силы

70 миллиграмм на одну линзу. Кроме массы перемещаемой жидкости на временные параметры гидролинзы влияют вязкость жидкости и конструкция гидравлической системы. Расчёт этой системы сводится к решению дифференциального уравнения второго порядка [35] и оптимизации значений параметров системы для удобства управления оптической силой гидролинзы.

3.2. Разрешение дисплея

Определим теперь угловой размер пикселя наголовного дисплея, показанного на рисунке 46. Используемый смартфон имеет экран размером $66 \times 145 \text{ мм}$ при разрешении 1080×2220 пикселей. При наблюдении с расстояния $f_1 = 0,046 \text{ м}$, согласно схеме рисунка 48, при $h = 66/2 = 33 \text{ мм}$ тангенс половины полного угла наблюдения по вертикали $\text{tg}\theta = h/f_1 = 0,717$, что соответствует $2\theta = 71,3$ градуса, а угловой размер пикселя $\Delta\theta = 2\theta/1080 \approx 4'$. Заметим, что это в четыре раза больше углового разрешения глаза, что годится для игрушки, но слишком много для профессиональных целей.

Управление оптической силой гидролинзы в наголовном дисплее виртуальной реальности необходимо производить для согласования аккомодации и конвергенции глаз пользователя при наблюдении объектов виртуального мира, находящихся на различной дистанции. Информацию о расстоянии до наблюдаемого объекта виртуального мира нельзя получить, направляя оптический локатор (лидар, *ToF*) на объект наблюдения, как в физическом мире,

например, как в описанных выше адаптивных очках. Виртуальные объекты не подлежат светолокации, однако расстояние до них можно получить методом триангуляции, отслеживая направление взгляда глаз пользователя стереоскопического дисплея.

3.3. Отслеживание глаз

Отслеживание глаз (*eye tracking*) можно осуществлять несколькими методами. Один из методов использует видеотрекеры. Большинство современных трекеров освещают глаз ближним инфракрасным неколлимированным (расходящимся) пучком света для создания роговичного отражения (блика). Вектор между центром зрачка и роговичным отражением можно использовать для вычисления направления взгляда. На рисунке 53 а красным перекрестьем обозначен центр зрачка, а зелёным – центр роговичного отражения и вектор между ними. На рисунке 53 б показан вид глаза при инфракрасной подсветке, а на рисунке 54 показаны расположения ИК светодиода и ТВ камеры наголовного дисплея виртуальной реальности для отслеживания положения правого глаза.

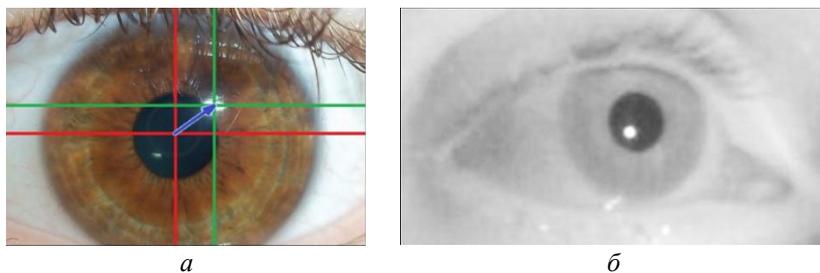


Рисунок 53 – Вектор между центром зрачка и роговичным отражением – *a*; инфракрасная подсветка глаза – *б*

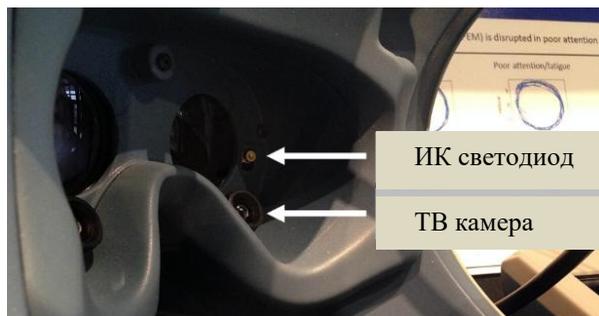


Рисунок 54 – Отслеживание глаз (*eye tracking*) в наголовном дисплее

Аналогичные светодиод и камера расположены с левой стороны дисплея. Два вектора для левого и правого глаз позволяют определить направление и дистанцию наблюдения R в виртуальном пространстве по каждому кадру изображений глаз. Поскольку физиология зрения такова, что вначале человек сводит (конвергирует) глаза на точку своего внимания в пространстве, а затем с опозданием на 0,25 секунды фокусирует (аккомодирует) хрусталики глаз, появляется возможность своевременно выставить нужную дистанцию R наблюдения для правильной аккомодации, регулируя оптическую силу гидролинзы D_{Γ} , в соответствии с полученным из (17) выражением

$$D_{\Gamma} = \frac{R - R_{\min}}{R \cdot R_{\min}}. \quad (19)$$

Таким образом, стереоскопический наголовный дисплей, оснащённый гидролинзами, позволяет обеспечить согласованную работу механизмов конвергенции и аккомодации глаз пользователя. Когда пользователь направляет своё внимание на определённый объект виртуального мира, его опыт оценки расстояний до объектов физического мира позволяет правильно конвергировать (сводить) глаза на объект для минимизации диспаратности сетчаточных изображений, а вслед за этим дисплей обеспечивает, корректируя оптическую силу гидролинз, правильную аккомодацию (фокусировку) глаз на определённую по конвергенции дистанцию. При этом, однако, при правильной фокусировке на объект наблюдения, остальные объекты, наблюдаемые боковым зрением, не получают необходимой правильной расфокусировки, поскольку кривизна фронта излучения от всех точек стереопары одинакова.

3.4. Размытие вне фокуса

Для полной иллюзии подобия обзора виртуального и физического миров необходимо обеспечить дополнительное размытие (блюр) изображения объектов, находящихся ближе или дальше выбранной дистанции фокусировки. Заметим, что в соответствии с рисунком 19 сетчатка глаза имеет существенно неравномерную разрешающую способность, которая обеспечивает размытие изображения вне узкой зоны ясного зрения (фовеа сетчатки размером около 2 градусов), что маскирует излишнюю чёткость бокового зрения, но не полностью устраняет её.

3.5. Динамический фовеальный рендеринг

Наличие неравномерной разрешающей способности глаза не только может играть положительную роль в решении задачи согласования конвергенции и аккомодации, но и позволяет значительно сократить информационный поток при обработке визуальной информации. Соответствующая технология получила название «динамический фовеальный рендеринг» (DFR). Одной из ведущих компаний, использующих эту технологию, является *Tobii Technology, Inc.* [36]. На рисунке 55 приведена фотография изображения на мониторе с DFR при отслеживании взгляда пользователя. На ней выделены три зоны, соответствующие зонам сетчатки глаз: фовеальная, средняя и периферийная. Технология работает особенно эффективно при высоком разрешении монитора вплоть до изображений 8 К и используется в гарнитурах виртуальной реальности и очках дополненной реальности. В этих применениях технология позволяет, например, повысить частоту кадров изображения, что особенно важно при поворотах головы пользователя для исключения скачков изображения.



Рисунок 55 – Чёткость изображения монитора с динамическим фовеальным рендерингом (DFR) при отслеживании взгляда пользователя; выделены три зоны, соответствующие зонам сетчатки глаз: фовеальная, средняя и периферийная

При разработке систем с переменной разрешающей способностью особое значение имеют методы определения разрешающей способности телевизионной системы с помощью специальных испытательных изображений. Рассмотрим подробно одно из них, которое, кроме определения разрешающей способности телевизионной системы и её элементов, позволяет создать галерею «муаровых портретов» массивов цветных светофильтров различных фотодатчиков.

4. Испытательное изображение – кольца Френеля

При цифровой обработке изображений для тестирования результатов обработки широко используется испытательное изображение в виде концентрических колец (зон) Френеля [11], [37]. Кольца Френеля являются пространственным аналогом временного сигнала линейной частотной модуляции (ЛЧМ) и могут использоваться подобно сигналу ЛЧМ для определения передаточных характеристик систем передачи изображений. В оптике изображение колец Френеля, нанесённое на стеклянную пластину, может использоваться для совершенно другой цели – фокусировки света, подобно стеклянной линзе. Кольца (зоны) Френеля названы так в честь Огюста Жана Френеля – французского физика, одного из создателей волновой теории света. Радиальный профиль яркости колец Френеля $F(r)$ может быть бинарным – двухступенчатым или синусоидальным. На рисунке 56 приведён вид колец Френеля с бинарным и синусоидальным профилями.

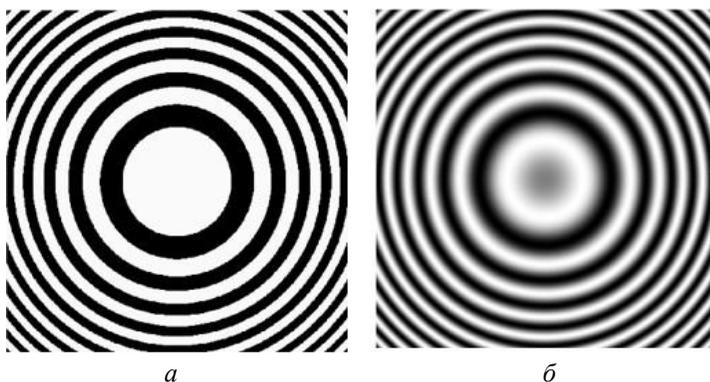


Рисунок 56 – Кольца Френеля с радиальными профилями: бинарным – *a*; синусоидальным – *б*

Все чёрные и белые кольца имеют примерно одинаковую площадь поверхности. В испытательном изображении радиусы границ между светлыми и тёмными кольцами Френеля определяются выражением

$$r_k = r_1 \sqrt{k} \quad (20)$$

где r_1 – радиус первого кольца (центрального круга) для бинарного профиля и первой от центра окружности с нулевым значением функции синус для синусоидального профиля.

Выберем значения двух уровней яркости для бинарных колец Френеля: 1 – для максимального, 0 – для минимального. При этом радиальный позитивный (p) (при светлом центральном круге) профиль яркости бинарных (b) колец Френеля можно определить выражением:

$$F^{bp}(r) = 1 - \left\lfloor \frac{r^2}{r_1^2} \right\rfloor + 2 \left\lfloor \frac{r^2}{2r_1^2} \right\rfloor, \quad (21)$$

где $\lfloor \cdot \rfloor$ – обозначение Айверсона функции «пол» (*floor*) – округление вниз; r – значение радиальной координаты изображения колец.

Такое задание функции профиля обеспечивает быстрое её вычисление. При произвольных уровнях яркости: максимальном U_{\max} и минимальном U_{\min} , – заданный профиль яркости колец Френеля определяется выражением:

$$U(r) = U_{\min} + (U_{\max} - U_{\min}) \cdot F(r). \quad (22)$$

Инверсный – негативный (n) (при тёмном центральном кольце) профиль яркости бинарных колец Френеля определяется выражением:

$$F^{bn}(r) = \left\lfloor \frac{r^2}{r_1^2} \right\rfloor - 2 \left\lfloor \frac{r^2}{2r_1^2} \right\rfloor. \quad (23)$$

Графики выражений (21) и (23) приведены на рисунке 57 пунктирными линиями.

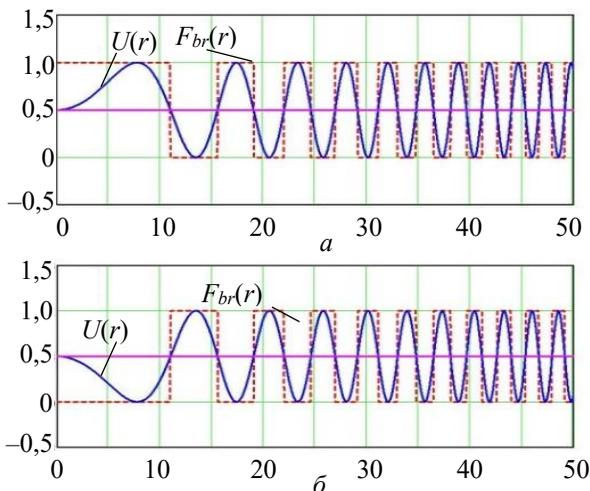


Рисунок 57 – Профили бинарных (пунктирные линии) и синусоидальных (сплошные линии) колец Френеля со светлым (a) и тёмным (b) центрами

Соответствующие радиальные синусоидальные профили яркости колец Френеля определяются выражениями

$$F^{sp}(r) = \frac{1}{2} \left[1 + \sin \left(\frac{\pi \cdot r^2}{r_1^2} \right) \right], \quad (24)$$

$$F^{sn}(r) = \frac{1}{2} \left[1 - \sin \left(\frac{\pi \cdot r^2}{r_1^2} \right) \right], \quad (25)$$

графики которых приведены на рисунке 57 сплошными линиями.

Аргумент функции \sin выражений (24) и (25) можно представить в виде:

$$\pi \cdot r^2 / r_1^2 = 2\pi \cdot v(r) \cdot r, \quad (26)$$

где $v(r) = r / 2r_1^2$ – пространственная частота, линейно зависящая от радиальной координаты r . Аналогичная зависимость от временной координаты при частотной модуляции свип-генератора, используемого совместно с осциллографом, применяется при анализе радиотехнических цепей с целью получения их амплитудно-частотных характеристик (АЧХ).

Систему передачи изображения можно оценивать формой её контрастно-частотной (частотно-контрастной) характеристики КЧХ (ЧКХ), которая определяется как отношение величины контраста репродукции (изображения, получаемого с помощью воспроизводящей системы) к контрасту соответствующей области оригинала (воспроизводимого объекта). КЧХ системы, воспроизводящей изображение, в общем случае зависит от направления измерения пространственной частоты из-за возможной анизотропии системы. Примером анизотропной системы является цифровая телевизионная система. Удобство использования колец Френеля в качестве оригинала обусловлено центральной симметрией их изображения, что обеспечивает необходимую изотропию оригинала.

Выражениями (21)–(25) с помощью алгебраических символов мы задаём непрерывный (континуальный) образ колец Френеля. Изображение колец на выходе телекамеры дискретно, то есть состоит из отдельных дискретов – элементов – пикселей (*picture's element – pix element – pixel*). Можно выделить два варианта преобразования континуального образа колец в их цифровое изображение:

- с выборкой континуального значения в центре пикселя;
- с усреднением континуальных значений по площади пикселя.

Первый вариант может использоваться для создания эталонного (оригинального) испытательного изображения с помощью принтера. Второй вариант моделирует создание репродукции оригинала с помощью цифровой фотокамеры. Рассмотрим далее оба варианта для синусоидального и бинарного профилей колец.

4.1. Кольца Френеля с синусоидальным профилем

Введём прямоугольную систему координат с осью абсцисс x и ординат y . При этом изображение колец Френеля с синусоидальным профилем и со светлой центральной зоной в соответствии с выражениями (24) и (25) будет определяться как

$$F^s(x, y) = \frac{1}{2} \left[1 \pm \sin \left(\frac{\pi \cdot (x^2 + y^2)}{r_1^2} \right) \right], \quad (27)$$

поскольку справедливо равенство $r^2 = x^2 + y^2$, и где знак плюс перед функцией $\sin(x)$ соответствует позитивному, а минус – негативному изображению колец Френеля.

Представим цифровое изображение колец набором квадратных пикселей с центрами в узлах регулярной прямоугольной решетки, наложенной на континуальное изображение колец. Яркость пикселей определяется яркостью континуального изображения колец в узлах решетки. Такое изображение можно определить прямоугольной матрицей (массивом) \hat{F} . Введём нумерацию элементов цифрового изображения начиная с его верхнего левого угла; при этом номер строки матрицы $i = 0, 1, 2, \dots, M - 1$, номер столбца $j = 0, 1, 2, \dots, N - 1$, а размер матрицы $M \times N$. Элемент изображения (матрицы, массива) обозначим символом $\hat{F}_{i,j}$. Пусть начало координат x и y (центр колец Френеля) совпадает с центром пикселя $i0, j0$, тогда в выражении (8) $x = j - j0$, $y = i0 - i$. Получим выражение для значений яркости элементов цифрового позитивного изображения колец с синусоидальным профилем с выборкой континуального значения в центре пикселя

$$\hat{F}_{i,j}^{sp} = \frac{1}{2} \left[1 + \sin \left(\frac{\pi \cdot ((j - j0)^2 + (i0 - i)^2)}{r_1^2} \right) \right]. \quad (28)$$

Аналогично, из (27), получим выражение для значений яркости элементов цифрового негативного изображения колец с синусоидальным профилем

$$\dot{F}_{i,j}^{sn} = \frac{1}{2} \left[1 - \sin \left(\frac{\pi \cdot ((j - j_0)^2 + (i_0 - i)^2)}{r_1^2} \right) \right]. \quad (29)$$

Выражения (28) и (29) позволяют синтезировать растровое (цифровое) изображение колец Френеля с синусоидальным профилем при использовании, например, пакета *Mathcad*. Проведём усреднение значений континуального образа колец по площади каждого из квадратных пикселей, на которые мы разбиваем этот образ. Такое усреднение можно представить интегрированием выражения (27) в пределах размеров пикселя $\Delta \times \Delta$ с последующей выборкой значений интеграла в центрах пикселей. Запишем выражение для интеграла:

$$\begin{aligned} \bar{F}^s(x, y) &= \int_{y-\frac{\Delta}{2}}^{y+\frac{\Delta}{2}} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} F^s(x, y) dx dy = \\ &= \frac{1}{2} \int_{y-\frac{\Delta}{2}}^{y+\frac{\Delta}{2}} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \left[1 \pm \sin \left(\frac{\pi \cdot (x^2 + y^2)}{r_1^2} \right) \right] dx dy, \end{aligned} \quad (30)$$

где знак плюс перед функцией $\sin(x)$ соответствует позитивному, а минус – негативному изображению колец Френеля.

Интеграл в правой части выражения (28) сводится к интегралам Френеля, которые не могут быть выражены в элементарных функциях. Вместе с тем, с хорошим приближением:

$$\int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \sin(ax^2) dx \approx \frac{\sin(\Delta ax)}{ax} \sin(ax^2) = \Delta \text{sinc}(\Delta ax) \sin(ax^2), \quad (31)$$

$$\int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \cos(ax^2) dx \approx \frac{\sin(\Delta ax)}{ax} \cos(ax^2) = \Delta \text{sinc}(\Delta ax) \cos(ax^2). \quad (32)$$

С учётом (31), (32) и в соответствии с (30) изображение колец Френеля с синусоидальным профилем, определяется выражением:

$$\begin{aligned} \bar{F}^s(x, y) \approx \tilde{F}^s(x, y) = \Delta^2 \left\{ 1 \pm \operatorname{sinc} \left(\frac{\pi \cdot \Delta \cdot x}{r_1^2} \right) \operatorname{sinc} \left(\frac{\pi \cdot \Delta \cdot y}{r_1^2} \right) \times \right. \\ \left. \times \sin \left[\frac{\pi}{r_1^2} (x^2 + y^2) \right] \right\}. \end{aligned} \quad (33)$$

Предполагая, что расстояние между центрами пикселей, расположенными в узлах прямоугольной решетки, равно единице, а начало координат x и y (центр колец Френеля) совпадает с центром пикселя $i0, j0$, получим выражение для значений яркости элементов (пикселей) цифрового изображения со светлым (знак плюс перед sinc функцией) или темным (знак минус перед sinc функцией) центром:

$$\begin{aligned} \bar{F}_{i,j}^s \approx \tilde{F}_{i,j}^s = \Delta^2 \left\{ 1 \pm \operatorname{sinc} \left[\frac{\pi \cdot \Delta}{r_1^2} (j - j0) \right] \operatorname{sinc} \left[\frac{\pi \cdot \Delta}{r_1^2} (i0 - i) \right] \times \right. \\ \left. \times \sin \left[\frac{\pi}{r_1^2} \left((j - j0)^2 + (i0 - i)^2 \right) \right] \right\}. \end{aligned} \quad (34)$$

где $\Delta \leq 1$.

Это выражение позволяет синтезировать растровое изображение колец Френеля с синусоидальным профилем с усреднением значений континуального образа колец по площади каждого из квадратных пикселей при использовании, например, пакета *Mathcad* и таким образом моделировать работу цифровой телевизионной камеры. На рисунке 58 а приведено цифровое изображение колец Френеля, синтезированное в соответствии с выражением (28) с учётом соотношения (22). Кроме основного центрального изображения колец это изображение содержит множество его прямых (позитивных) и инверсных (негативных) копий, так называемых посторонних узоров (ПУ). Наличие ПУ на изображении колец Френеля является характерным признаком цифрового изображения с регулярным растром. Используя выражения (28), (29) или (34) можно изучать ПУ, воспроизводя графики профилей их изображений, фиксируя какой-либо столбец или строку изображения. На рисунке 58 б приведено цифровое изображение колец Френеля, синтезированное в соответствии с выражением (34) с учётом соотношения (22).

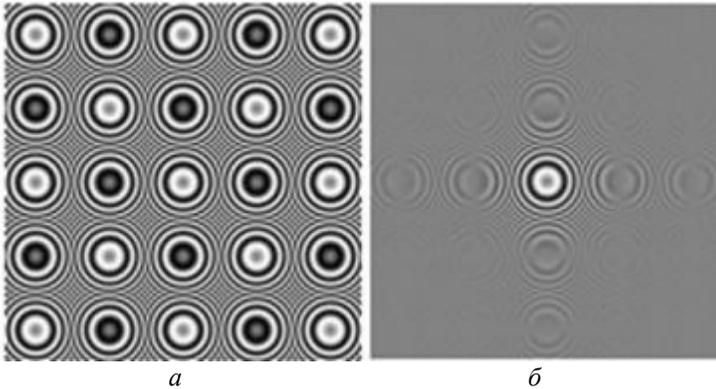


Рисунок 58 – Синтезированные изображения колец Френеля: прямой выборкой значений континуального образа колец в узлах сетки прямоугольного раstra – *a*; выборкой с предварительным усреднением по площади каждого пикселя – *б*

Кроме центрального изображения колец рисунок 58 *б* также содержит посторонние узоры, но они менее контрастные, чем основное центральное изображение колец, и их контраст убывает к периферии рисунка, что обусловлено модулирующим действием sinc-функций.

Изображения рисунка 58 синтезированы при значениях: $r_1 = 5$; $M = N = 120$; $i_0 = j_0 = 60$; $U_{\min} = 10$; $U_{\max} = 250$; $\Delta = 1$.

Поскольку выражение (34) является приближённым равенством, оно позволяет воспроизводить цифровое изображение колец Френеля с некоторой ошибкой. Ошибку синтеза изображения можно определить матрицей разностей элементов двух изображений, значение яркости которых вычисляются либо с использованием выражения (34), либо путём интегрирования выражения (30) с использованием численных методов. Разность элементов двух таких изображений определяется выражением:

$$\Theta_{i,j} = \tilde{F}_{i,j}^s - \bar{F}_{i,j}^s, \quad (35)$$

Матрицу таких элементов можно представить также в виде изображения. На рисунке 59 отображены значения увеличенной в 50 раз ошибки представления элементов колец Френеля с помощью выражения (35). Эта ошибка уменьшается к краям изображения, где сравнима с величиной одного уровня квантования. Напомним, что с учётом контрастной чувствительности зрения, цифровое изображение обычно квантуется на 256 уровней яркости.

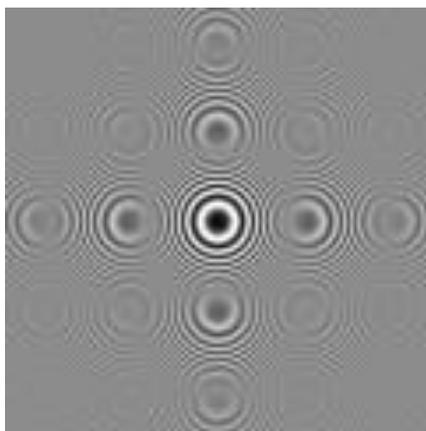


Рисунок 59 – Отображение усиленной в 50 раз ошибки

На рисунке 60 *a* приведён график изменения абсолютного значения этой ошибки вдоль шестидесятой (центральной) строки изображения колец, рассчитанный по формуле:

$$\Theta_{i0,j} = \tilde{F}_{i0,j}^s - \bar{F}_{i0,j}^s \quad (36)$$

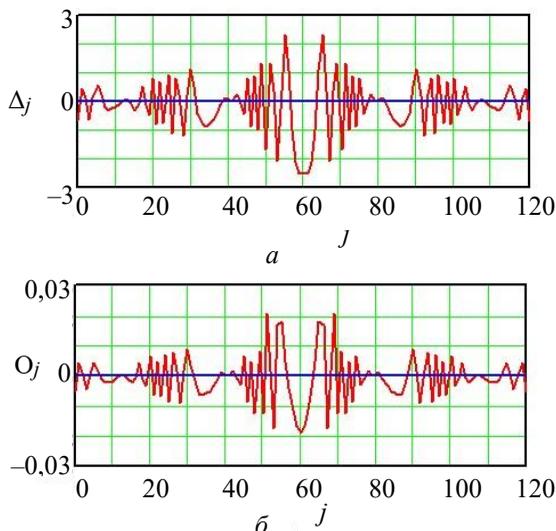


Рисунок 60 – Значение абсолютной (*a*) и относительной (*б*) ошибок представления колец Френеля с использованием выражения (34) вдоль строки изображения колец Френеля с номером $i = i_0$

Можно видеть, что эта ошибка не превышает значения трёх уровней квантования яркости. На рисунке 60 *б* приведён график

изменения относительного значения этой ошибки вдоль шестидесятой (центральной) строки изображения колец, рассчитанный по формуле

$$O_j = \frac{\Theta_{i0,j}}{\bar{F}_{i0,j}^s} \quad (37)$$

В данном случае относительная ошибка в основном меньше двух процентов, то есть зрительно практически не воспринимается, что оправдывает использование выражения (34).

4.2. Кольца Френеля с бинарным профилем

В прямоугольной системе координат континуальное изображение колец Френеля с бинарным профилем и светлой центральной зоной в соответствии с выражением (21) определяется выражением

$$F^{bp}(x, y) = 1 - \frac{x^2 + y^2}{r_1^2} + 2 \frac{x^2 + y^2}{2r_1^2} \quad (38)$$

Выражение для значений яркости элементов цифрового позитивного изображения колец с бинарным профилем с выборкой континуального значения в центре пикселя имеет вид:

$$\dot{F}_{i,j}^{bp} = 1 - \left[\frac{(j - j0)^2 + (i0 - i)^2}{r_1^2} \right] + 2 \left[\frac{(j - j0)^2 + (i0 - i)^2}{2r_1^2} \right] \quad (39)$$

Аналогично, из (23) получим выражение для значений яркости элементов цифрового негативного изображения колец с бинарным профилем

$$\dot{F}_{i,j}^{bn} = \left[\frac{(j - j0)^2 + (i0 - i)^2}{r_1^2} \right] - 2 \left[\frac{(j - j0)^2 + (i0 - i)^2}{2r_1^2} \right] \quad (40)$$

Выражения (39) и (40) позволяют синтезировать растровое изображение колец Френеля с бинарным профилем при использовании пакета *Mathcad*. На рисунке 61 приведено синтезированное изображение колец Френеля с бинарным профилем путём прямой выборки значений континуального образа колец в узлах сетки прямоугольного раstra в соответствии с выражением (36) с учётом соотношения (22) при значениях: $r_1 = 5$; $M = N = 120$; $i0 = j0 = 60$; $U_{\min} = 10$; $U_{\max} = 250$.

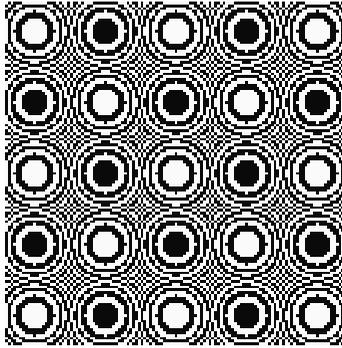


Рисунок 61 – Синтезированное изображение колец Френеля

Проведём усреднение значений континуального образа колец по площади каждого из квадратных пикселей, на которые мы разбиваем этот образ. Такое усреднение можно представить интегрированием выражения (38) в пределах размеров пикселя $\Delta \times \Delta$ с последующей выборкой значений интеграла в центрах пикселей. Проинтегрируем выражение (38):

$$\bar{F}^p(x, y) = \int_{y-\frac{\Delta}{2}}^{y+\frac{\Delta}{2}} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} F^p(x, y) dx dy = \int_{y-\frac{\Delta}{2}}^{y+\frac{\Delta}{2}} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \left(1 - \left[\frac{x^2 + y^2}{r_1^2} \right] + 2 \left[\frac{x^2 + y^2}{2r_1^2} \right] \right) dx dy. \quad (41)$$

Для интегрирования выражения (41) воспользуемся равенствами:

$$\int_0^x \left[\frac{x^2}{a^2} \right] dx = \frac{x^2}{2} \left[\frac{x^2}{a^2} \right] - \frac{a^2}{2} \sum_{k=0}^{\lfloor \frac{x^2}{a^2} \rfloor} k, \quad (42)$$

$$\int_{\frac{1}{a}}^x \frac{\lfloor x \rfloor}{\sqrt{ax-1}} dx = \frac{2}{a} \left[\lfloor x \rfloor \sqrt{ax-1} - \sum_{k=0}^{\lfloor x \rfloor} \text{Re}(\sqrt{ax-1}) \right], \quad (43)$$

$$\int_{\frac{1}{a}}^x \frac{\sum_{k=0}^{\lfloor x \rfloor} k}{x \sqrt{ax-1}} dx = 2 \left\{ \text{arctg} \left[\text{Re}(\sqrt{ax-1}) \right] \sum_{k=0}^{\lfloor x \rfloor} k - \right. \quad (44)$$

$$\left. - \sum_{k=0}^{\lfloor x \rfloor} k \cdot \text{arctg} \left[\text{Re}(\sqrt{ak-1}) \right] \right\}$$

Интегрирование удобно начинать в полярной системе координат r, φ в связи с центральной симметрией колец Френеля с последующим переходом к декартовой системе координат x, y при соотношениях между системами в соответствии с рисунком 62.

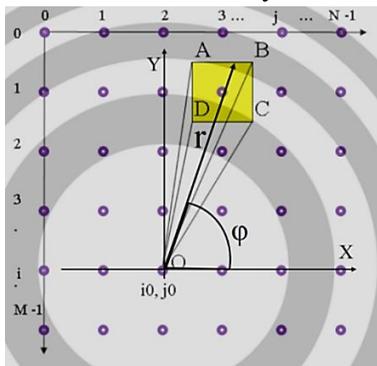


Рисунок 62 – Кольца Френеля в системах координат r, φ, X, Y и i, j

На этом рисунке также приведена система координат цифрового (растрового) изображения i, j с началом в верхнем левом углу изображения. Центры пикселей цифрового изображения обозначены небольшими колечками, а один из пикселей – квадратом с углами A, B, C, D . Начало O координат X, Y совпадает с центром одного из пикселей с координатами i_0, j_0 и с центром бинарных колец Френеля, показанных здесь в серых тонах.

Процедура интегрирования выражения (41) сводится к интегрированию по площадям треугольников OAB, OBC, ODC и OAD с последующим вычислением разности суммы выражений для двух первых треугольников и суммы выражений для двух последних. Обычно решения подобных задач на интегрирование имеют довольно громоздкий характер. Наш случай не исключение.

Листинг программы вычисления значений пикселей цифрового изображения колец Френеля с бинарным профилем в системе *Mathcad*, обеспечивающей максимальное приближение к естественной форме записи математических выражений, занимает две страницы текста без комментариев. Практически важно то, что расчёт по такой программе вполне укладывается в реальный масштаб времени. Заметим, что полученное нами решение для значения интеграла выражения (41) является точным, как и приведённые выражения (42)–(44).

На рисунке 63 приведено изображение колец Френеля с бинарным профилем, синтезированное с использованием полученного решения для значения интеграла выражения (41) с усреднением значений континуального образа колец по площади каждого квадратного пикселя изображения. Интересно сравнить этот рисунок с рисунками 61 и 58. Как и в случае перехода от выражения (33)

к выражению (34), рисунок 63 получен в предположении, что расстояние между центрами пикселей, расположенных в узлах прямоугольной решетки, равно единице, а начало координат x и y (центр колец Френеля) совпадает с центром пикселя i_0, j_0 . Изображение рисунка 63 позитивное, а уровни его яркости ограничены значениями $U_{\min} = 10$ и $U_{\max} = 250$, заданными при использовании зависимости, аналогичной используемой в выражении (41).

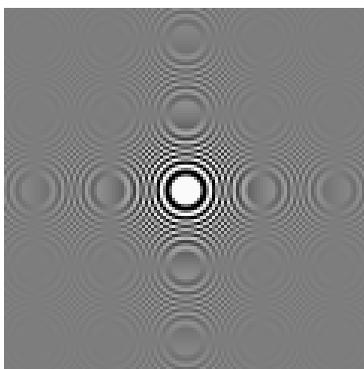


Рисунок 63 – Синтезированное изображение колец Френеля с бинарным профилем путем усреднения значений континуального образа колец по площади каждого пикселя изображения

Обычно, при съёмке испытательного изображения колец Френеля с помощью цифровой фото- или телекамеры центр колец не совпадает с центром пикселя i_0, j_0 . И поэтому картина посторонних узоров может отличаться от изображенной на приведённых выше рисунках. Для моделирования относительных сдвигов раstra цифрового изображения и колец Френеля следует ввести в приведённые выше математические выражения специальные параметры сдвига δ_x и δ_y .

4.3. Сдвиги колец Френеля относительно раstra изображения

Вводя параметры δ_x и δ_y , например, в выражение (34), различая горизонтальный и вертикальный размеры сенселя Δ_x и Δ_y и определяя постоянную составляющую F_0 и амплитуду F_1 синусоидального профиля, получим следующее выражение для элемента изображения колец Френеля:

$$\tilde{F}_{i,j}^s = \Delta_x \Delta_y \left\{ F_0 \pm F_1 \operatorname{sinc} \left[\frac{\pi \Delta_x}{r_1^2} (j - j_0 - \delta_x) \right] \operatorname{sinc} \left[\frac{\pi \Delta_y}{r_1^2} (i_0 - i - \delta_y) \right] \operatorname{sinc} \left[\frac{\pi}{r_1^2} (j - j_0 - \delta_x)^2 + (i_0 - i - \delta_y)^2 \right] \right\}. \quad (45)$$

Обратим внимание на то, что в выражении (45) параметры δ_x и δ_y вводятся со знаком минус в выражения для координат по осям x и y в соответствии с индексами параметров. При этом положительные значения самих параметров соответствуют сдвигу изображения колец вправо по оси x и вверх по оси y относительно системы координат x, y и привязанной к ней системы координат j, i . Отрицательные значения параметров δ_x и δ_y соответствуют сдвигу изображения колец в противоположных направлениях.

Поскольку мы произвольно выбираем (задаём) сенсель i_0, j_0 , в зону которого попадает центр колец Френеля, следует ограничить области задания параметров δ_x и δ_y так, чтобы при их введении центр колец не покидал зону заданного сенселя. Для неперекрывающихся сенселей значения параметров сдвига не превышают половины максимально возможных размеров сенселя (расстояния между сенселями) по горизонтали или вертикали, то есть:

$$0,5\Delta_x < \delta_x < 0,5\Delta_x, \quad 0,5\Delta_y < \delta_y < 0,5\Delta_y.$$

При сдвиге изображения колец изменяется радиус колец побочных узоров (ПУ). На рисунке 64 приведён набор из 12 картин колец ПУ первого порядка при последовательном сдвиге колец Френеля с интервалом $0,1\Delta_x$, при $\Delta_x = 1$. Радиусы полосатых колец ПУ соответствуют узким границам между кольцами, которые представляют собой зоны минимального контраста (различия) между сигналами соседних сенселей фотодатчика. Между этими границами контраст полосок колец ПУ нечётных порядков возрастает. Радиус кольца ПУ определяется выражением:

$$R_k = r_1 \sqrt{\left\{ \frac{r_1^2}{4} \left(\frac{m^2}{\Delta_x^2} + \frac{n^2}{\Delta_y^2} \right) + \frac{m\delta_x}{\Delta_x} + \frac{n\delta_y}{\Delta_y} \right\} + k}, \quad (46)$$

где фигурными скобками обозначена дробная часть числа $\{x\} = x - \lfloor x \rfloor$, $m = 0, \pm 1, \pm 2, \dots$ – порядок ПУ по оси x ; $n = 0, \pm 1, \pm 2, \dots$ – порядок ПУ по оси y ; $k = 0, 1, 2, \dots$ – номер кольца.

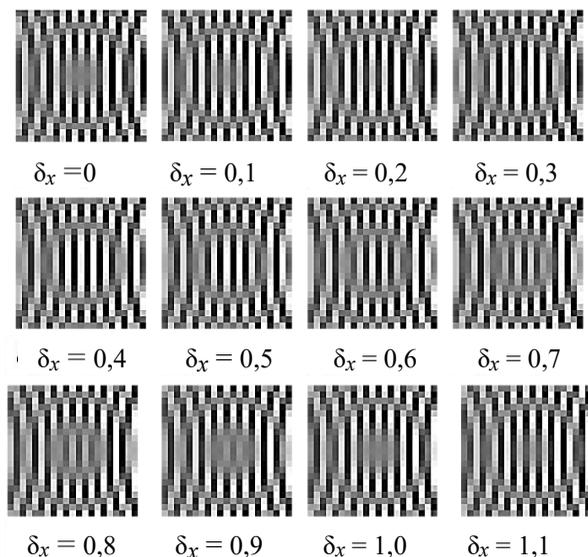


Рисунок 64 – Изменение радиуса колец постороннего узора (муара) первого порядка при изменении сдвига колец Френеля δ_x относительно начала координат вдоль оси x

Положение центра кольца постороннего узора, заданное координатами X_m и Y_n , определяется выражениями¹:

$$X_m = \frac{mr_1^2}{2\Delta_x} + \delta_x, \quad Y_n = \frac{nr_1^2}{2\Delta_y} + \delta_y. \quad (47)$$

С помощью выражений (47) можно выбирать положение центра подматрицы матрицы изображения колец Френеля, на которой отображается небольшая часть изображения, включающая зону ПУ, как это сделано для фрагментов (см. рисунок 64). Непосредственному наблюдению колец ПУ высших нечётных порядков мешает снижение контраста изображения соседних (полосатых) колец из-за ослабляющего действия sinc-функций выражения (45). Кольца ПУ чётных порядков, не имеющие полосатой структуры, испытывают инверсию и также теряют межкольцевой контраст с увеличением их порядка. Можно искусственно повысить контраст участка (подматрицы) изображения колец Френеля, содержащего зону ПУ заданного порядка. На рисунке 65,

¹ Вариант вывода выражений (46) и (47) для одномерной дискретизации приведён в работе [38] с использованием собственных обозначений.

полученном с использованием выражения (45), проведено местное контрастирование ПУ четвертого порядка ($m = -4$) по горизонтальной оси и второго ($n = 2$) по вертикальной в условиях, когда $\delta_x = \delta_y = 0$. При этом хорошо видны последствия четырехкратной инверсии яркости колец ПУ в районе нулей sinc-функций выражения (45).

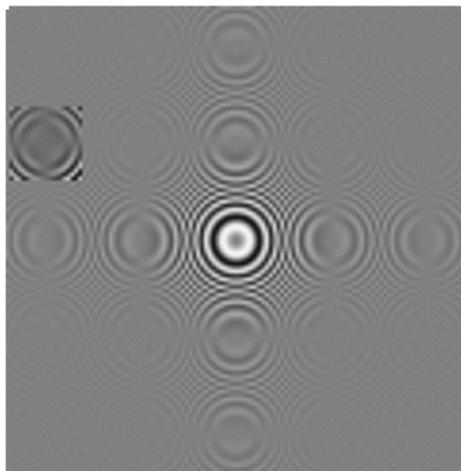


Рисунок 65 – Синтезированное изображение колец Френеля с применением местного контрастирования постороннего узора четвертого порядка

При математическом моделировании для устранения снижения контраста колец ПУ достаточно исключить в выражении (45) sinc-функции, тем самым переходя, по существу, к выражениям вида (28) или (29). Для случая использования изображения фотодатчика такая операция включает деление функции модуляции яркости этого изображения на соответствующие sinc-функции, что осложняется необходимостью деления на ноль.

В цифровой системе передачи изображений могут использоваться различные сетки раstra изображения. Как правило, это прямоугольные периодические структуры, хотя могут быть представлены и треугольные или шестиугольные (гексагональные) сетки. Подход к анализу таких структур при воспроизведении колец Френеля аналогичен подходу, использованному при анализе выражения (41).

Особый интерес вызывает применение колец Френеля для испытаний систем передачи цветного изображения, в которых

осуществляется цветоделение исходного оптического изображения. В случаях последовательного цветоделения, а также параллельного цветоделения с помощью полупрозрачных зеркал или призм, формируются одноцветные компоненты с кольцами Френеля, анализ которых ничем не отличается от вышеизложенного. В современных видеосистемах наиболее распространенным является метод пространственного цветоделения с использованием больших массивов цветных светофильтров. В качестве примера на рисунке 66 *a* приведены фрагменты различных массивов цветных светофильтров, осуществляющих пространственное цветоделение изображения. При моделировании воспроизведения чёрно-белых колец Френеля в таких системах передачи следует учитывать ряд особенностей одноцветных компонент. В одном случае, например при использовании фильтра Байера, каждая одноцветная (красная, зелёная или синяя) компонента формируется из сигналов массива сенселей, находящихся под светофильтрами данного цвета, путём интерполяции значений сигналов этих сенселей в местах соответствующих положениям сенселей других цветовых компонент. Массивы сенселей различных компонент отличаются здесь друг от друга *сдвигом*, кратным межсенсельному расстоянию фотодатчика относительно центра колец Френеля. В других случаях компоненты могут отличаться также периодами своих сеток, комбинациями периодов и содержанием сигналов соответствующего цвета в сигналах сенселей со светофильтрами комбинированных (неосновных) цветов. Во всех случаях при модельном синтезе цветного изображения колец Френеля следует для любого заданного массива цветных светофильтров (см. рисунок 66 *a*) определить сетки (матрицы) фильтров одного цвета с индексом Col с элементами $q_{i,j}^{Col}$, такими, что $q_{i,j}^{Col} = 1$, если светофильтр сенселя соответствует указанному цвету и $q_{i,j}^{Col} = 0$, если нет. Затем определить сигналы исходных изображений одноцветных компонент в соответствии с выражением:

$$Q_{i,j}^{Col} = q_{i,j}^{Col} F_{i,j}, \quad (48)$$

где $F_{i,j}$ – элемент чёрно-белого изображения колец Френеля, индексы которого i и j соответствуют выражению (28).

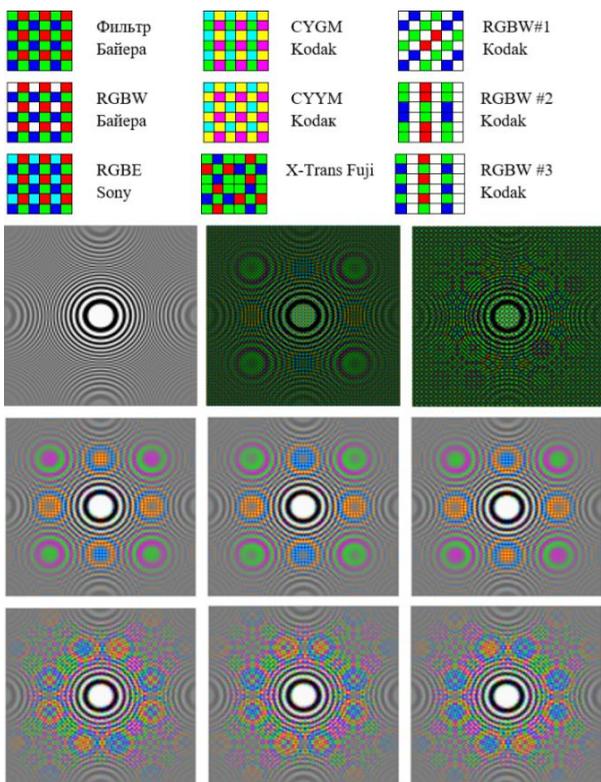


Рисунок 66 – Массивы цветных светофильтров различных фотодатчиков (а); цифровые изображения колец Френеля с прямоугольным (бинарным) радиальным профилем и усреднением по площади сенселя (б – к):

а		
б	в	г
д	е	ж
з	и	к

чёрно-белого фотодатчика – б; с фильтром Байера без интерполяции – в; с *X-Trans* фильтром без интерполяции – г; с фильтром Байера с интерполяцией и сдвигом влево на половину сенселя – д; с фильтром Байера без сдвига – е; с фильтром Байера со сдвигом вправо – ж; с *X-Trans* фильтром с интерполяцией и сдвигом влево – з; с *X-Trans* фильтром без сдвига – и; с *X-Trans* фильтром со сдвигом вправо – к

На всех изображениях рисунка 66 (б – к) хорошо видны посторонние узоры, связанные с дискретизацией изображения колец.

Существуют различные алгоритмы интерполяции значений цветовой компоненты в местах сенселей со светофильтрами другого цвета [39].

Определим сигналы изображений компонент основных цветов в случае линейной интерполяции для массивов основных цветов для случаев применения фильтров Байера и *X-Trans Fuji* (см. рисунок 66 а) в соответствии с выражением:

$$\text{Если } q_{i,j}^{Col} = 0, \text{ то } Q_{i,j}^{ICol} = \frac{\sum_{k=-1}^{k=1} \sum_{l=-1}^{l=1} Q_{i+k,j+l}^{Col}}{\sum_{k=-1}^{k=1} \sum_{l=-1}^{l=1} q_{i+k,j+l}}, \quad (49)$$

$$\text{если нет, то } Q_{i,j}^{ICol} = Q_{i,j}^{Col}.$$

где *col* индекс обозначения цвета; в соответствии с рисунком 66 а для индекса *Col* возможны обозначения: *R* – красный, *B* – синий, *G* – зелёный.

При такой интерполяции следует ввести граничные условия, например:

$$Q_{i,j}^{ICol} = \Gamma = 0, \text{ если: } i = j = 0, \text{ или } i = M, \text{ или } j = N. \quad (50)$$

В выражении (49) для интерполяции используется подматрица размером 3×3.

Синтезируем, например, цветное изображение колец Френеля с использованием массива фильтра Байера. Введём функцию – указатель кратности двух чисел *a* и *b*:

$$u(a,b) = 1 \frac{a}{b} + \frac{a}{b}, \quad (51)$$

где $\lceil \cdot \rceil$ – обозначение Айверсона функции «потолок» – округление вверх, при этом $u(a,b) = 1$ если *a* делится на *b* без остатка и $u(a,b) = 0$, если нет.

Сетки для красной (*red*), синей (*blue*) и зелёной (*green*) компонент цветного изображения задаются соответствующими выражениями:

$$\begin{aligned} r_{i,j}^B &= q_{i,j}^R = u(i,2)u(j+1,2), \\ b_{i,j}^B &= q_{i,j}^B = u(i+1,2)u(j,2), \\ g_{i,j}^B &= q_{i,j}^G = 1 - r_{i,j}^B - b_{i,j}^B, \end{aligned} \quad (52)$$

а сигналы исходных изображений одноцветных компонент: красной $R_{i,j}^B$, зелёной $G_{i,j}^B$ и синей $B_{i,j}^B$ изображения колец Френеля,

например, с синусоидальным профилем, задаваемым выражением (45), определяются в соответствии с выражениями:

$$\tilde{R}_{i,j}^B = Q_{i,j}^R = r_{i,j}^B \tilde{F}_{i,j}^s, \quad \tilde{G}_{i,j}^B = Q_{i,j}^G = g_{i,j}^B \tilde{F}_{i,j}^s, \quad \tilde{B}_{i,j}^B = Q_{i,j}^B = b_{i,j}^B \tilde{F}_{i,j}^s. \quad (53)$$

Сигналы интерполированных изображений одноцветных компонент

$$\tilde{R}_{i,j}^{BI} = Q_{i,j}^{IR}, \quad \tilde{G}_{i,j}^{BI} = Q_{i,j}^{IG}, \quad \tilde{B}_{i,j}^{BI} = Q_{i,j}^{IB}, \quad (54)$$

определяются в соответствии с выражением (49).

Описанные математические модели дискретизации испытательных изображений в виде колец Френеля позволяют рассчитать идеальные картины воспроизведения колец с помощью различных датчиков изображений. Такие картины необходимы при разработке методик и средств измерений для оценки качества систем передачи и обработки изображений. Обратим внимание на то, что посторонние узоры (муары) характерны для каждого типа цветных светофильтров (матрицы) и изменяются при сдвиге положения колец Френеля относительно растра датчика. Сдвигая положение колец с заданным небольшим шагом и последовательно фиксируя каждое полученное изображение, можно создать анимацию непрерывного преобразования – сдвига на величину целого пикселя, что позволяет создать галерею живых «муарных портретов» различных фотодатчиков (*gif* – файлов).

Фрагменты колец Френеля используются в качестве элементов стандартных телевизионных испытательных таблиц, например, в таблице ИТ-72, (СССР, 1972), предназначенной для испытаний аналоговых телевизионных систем при определении их разрешающей способности. Для испытаний цифровых теле- и фотосистем используются методы, принятые в международном стандарте ISO 12233 (2000, 2014, 2017...), в частности, метод наклонного края [40].

Как и в случае с динамическим фовеальным рендерингом (DFR), при отслеживании взгляда пользователя (см. рисунок 55) использование свойств зрения человека может помочь осуществить эффективное кодирование изображения (сжатие информации) без потери качества его восприятия пользователем. Опишем результаты моделирования некоторых свойств зрения, которые могут использоваться при разработке методов кодирования объёмного изображения.

5. Зрение

5.1. Контурное кодирование изображений

Кодирование изображений может преследовать различные цели, например, защиту визуальной информации от перехвата или сокращение её избыточности при передаче по каналам связи. При кодировании цифрового изображения производится его преобразование – фильтрация без потерь или с потерей некоторой части визуальной информации. Обычно потери допускаются с учётом того, что конечный приёмник информации (телезритель, робот) не ощущает этой потери, поскольку она находится за пределом его восприятия. В технических приложениях для сокращения избыточности информации используются различные методы сжатия (компрессии) изображений и число этих методов весьма велико [16]. Среди них, обычно, отсутствуют методы, связанные с выделением контуров объектов изображения. В системах технического зрения, наоборот, методы выделения контуров с целью улучшения восприятия или распознавания изображений используются широко [11], [41].

Другая ситуация наблюдается в биологии. Например, зрительная система человека устроена таким образом, что в зрительную кору головного мозга из каждого глаза поступает контурное изображение объектов наблюдения. Это обстоятельство обусловлено тем, что многие волокна зрительного нерва имеют центрально-симметричную *on-off* организацию своих рецептивных полей [6]. Математическое описание рецептивных полей дано, например, в [42]. В случае зрительной системы контурное изображение, по-видимому, позволяет ускорить реакцию организма на контур – форму объектов наблюдения, не отягощённую избыточной информацией об уровнях освещения их поверхностей. Кроме того, переход к контурам позволяет свести до минимума энергетические затраты организма в процессе передачи информации от глаза к зрительной коре, поскольку в этом случае средняя частота импульсов действия значительно снижается. В принципе, при известных значениях исходного изображения на его границах, контурное изображение может быть полностью восстановлено до исходного. Восстановление (декодирование) контурного изображения отличается простотой алгоритмов, что особенно важно при использовании методов параллельной обработки изображений для получения максимального эффекта от их использования.

5.2. Выделение контура

Контур объекта цифрового изображения может быть выделен с помощью свёртки изображения с некоторой маской (ядром, кернам, матрицей, оператором, весовой функцией). В зависимости от вида маски и содержания изображения контур может иметь различный профиль (поперечное сечение). Часто используются квадратные маски различных размеров: 2×2 , 3×3 , 5×5 элементов и др. На рисунке 67 приведен пример масок размером 3×3 элемента.

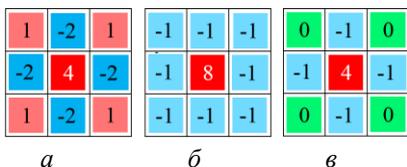


Рисунок 67 – Маски фильтров выделения контуров: оператор дискретного аналога двумерной производной второго порядка – *a*; однородно-периферийный оператор – *b*; дискретный аналог оператора Лапласа – *v*

Здесь цветом выделены положительные (оттенки красного), отрицательные (оттенки синего) и нулевые (оттенок зелёного) значения весов ячеек маски. Сумма весов каждой маски равна нулю. Маска на рисунке 67 *a* получена как результат последовательного применения прямой и обратной конечных разностей, как в вертикальном, так и в горизонтальном направлениях. Маска на рисунке 67 *b* получена заданием однородной периферии.

Маска на рисунке 67 *v* получена как сумма применений второй конечной разности в вертикальном и горизонтальном направлениях и по аналогии со случаем частных производных второго порядка называется оператором Лапласа.

Покажем вариант получения маски рисунка 67 *a*. Обозначим символом $v_{i,j}$ значение элемента изображения из i -той строки в вертикальном направлении и j -го столбца в горизонтальном направлении изображения, тогда обратная конечная разность элементов в вертикальном направлении определяется выражением:

$$f_{i,j} = v_{i,j} - v_{i-1,j}. \quad (55)$$

Прямая конечная разность элементов в вертикальном направлении:

$$g_{i,j} = f_{i+1,j} - f_{i,j}. \quad (56)$$

Вторая симметричная конечная разность элементов в вертикальном направлении определяется подстановкой выражения (55) в выражение (56):

$$g_{i,j} = v_{i-1,j} - 2v_{i,j} + v_{i+1,j}. \quad (57)$$

Обратная конечная разность элементов в горизонтальном направлении определяется выражением:

$$h_{i,j} = g_{i,j} - g_{i,j-1}. \quad (58)$$

Прямая конечная разность элементов в горизонтальном направлении:

$$e_{i,j} = h_{i,j+1} - h_{i,j}. \quad (59)$$

Вторая конечная симметричная разность элементов в горизонтальном направлении определяется подстановкой выражения (58) в выражение (59):

$$e_{i,j} = g_{i,j-1} - 2g_{i,j} + g_{i,j+1}. \quad (60)$$

Подстановкой выражения (58) в выражение (60) после приведения подобных членов и переноса начала отсчёта в правый нижний угол маски получим выражение для второй конечной разности:

$$E_{i,j} = v_{i-2,j-2} + v_{i-2,j} + v_{i,j-2} + v_{i,j} + 4v_{i-1,j-1} - 2v_{i-1,j-2} - 2v_{i-1,j} - 2v_{i-2,j-1} - 2v_{i,j-1}, \quad (61)$$

которое соответствует маске, изображенной на рисунке 67 а.

Нумерацию строк и столбцов изображения будем начинать с его левого верхнего угла с цифры «1» и заканчивать значениями: M для строк и N для столбцов (где M и N число строк и столбцов изображения, соответственно). Заметим, что в выражении (55) при вычислении элемента с начальным значением индекса $i = 1$ возникает трудность, поскольку неизвестно значение элемента с индексом $i - 1 = 0$. Аналогично в выражении (56) для конечного значения индекса $i = M$ требуется неизвестное значение элемента с индексом $i + 1 = M + 1$. Та же ситуация с расчётами по индексу j . Трудность преодолевается введением понятия о границе изображения и требования вести расчёт для неграничных элементов. При этом значения граничных элементов считаются известными. В выражении (61) будем считать границей две верхние строки и два левых столбца, а расчёт вести для элементов в интервалах номеров: $3 \leq i \leq M$ и $3 \leq j \leq N$.

Произведём свёртку тестового изображения v с маской рисунка 67 а, описываемую выражением (61). На рисунке 68 приведены изображения: тестовое v ; масштабированное E -контурное; контрастированное контурное с обрезкой модуля значений пикселей больше 120 и масштабированное (E); контурное в псевдоцветах с красными положительными и синими отрицательными значениями контуров (E). Масштабирование

в интервал 0–255 позволяет отобразить отрицательные значения контуров в виде их тёмных аналогов, а положительные – в светлых. На рисунке 68 б можно видеть, что масштабированное изображение состоит из пикселей, яркость которых в основном близка к среднему серому значению. Поэтому, для наглядности, на рисунке 68 в приведён его искусственно контрастированный (искажённый по яркости) аналог, а на рисунке 68 г положительные значения контуров изображены красным цветом, а отрицательные – инвертированы в положительные и изображены синим цветом. Отметим особенность формирования маски рисунке 67 а, которая состоит в том, что конечные разности вида (55), (56), (58), (59) можно брать в любом порядке, что не меняет вида выражения (61).

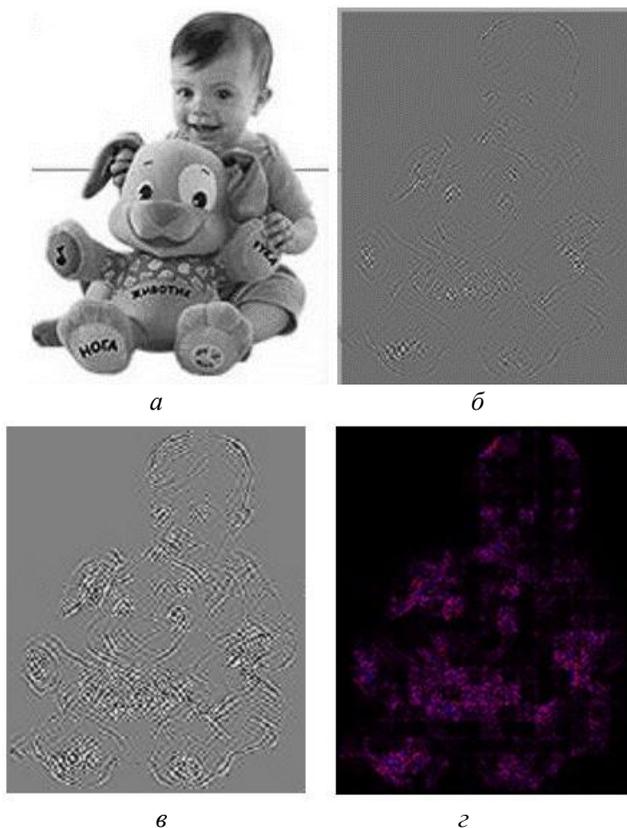


Рисунок 68 – Изображения: исходное $v - a$;
контурное масштабированное $E - б$;

искусственно контрастированное $E - в$; псевдо-цветное контурное $E - г$

Произведём свёртку тестового изображения v с маской рисунка 67 б, описываемую выражением:

$$K_{i,j} = 8v_{i-1,j-1} - v_{i-2,j-2} - v_{i-2,j} - v_{i,j-2} - v_{i,j} - v_{i-1,j-2} - v_{i-1,j} - v_{i-2,j-1} - v_{i,j-1} \quad (62)$$

На рисунке 69 приведены изображения: тестовое v ; масштабированное контурное K ; контрастированное контурное с обрезкой модуля значений пикселей больше 120 и масштабированное K ; контурное в псевдоцветах с красными положительными и синими отрицательными значениями контуров K . Все изображения на рисунке 69 получены аналогично изображениям на рисунке 68, но с другой маской.

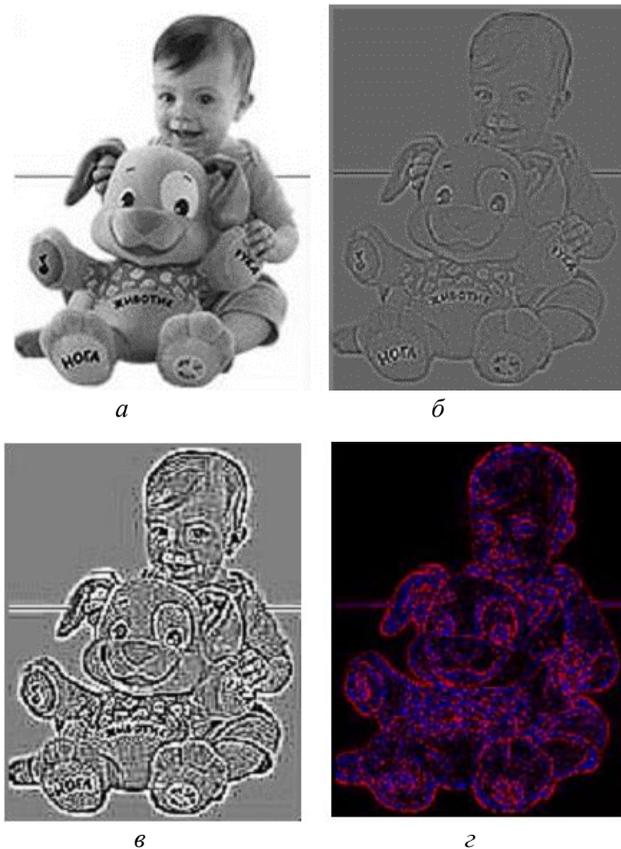


Рисунок 69 – Изображения: исходное $v - a$;
контурное масштабированное $K - б$;
искусственно контрастированное $K - в$; псевдо-цветное контурное $K - г$

Маски рисунка 67 *а, б* иногда называют операторами Лапласа из-за их центральной симметрии, чего явно недостаточно для такого названия. Маска на рисунке 67 *в* является дискретным аналогом оператора Лапласа:

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (63)$$

Это выражение является суммой коэффициентов матрицы (маски), полученной суммированием двух матриц:

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}. \quad (64)$$

где первое слагаемое соответствует выражению (57), а второе – выражению (60).

Математическое выражение для маски рисунка 67 *в* содержит всего 5 членов, поскольку угловые коэффициенты этой маски равны нулю, и имеет вид

$$L_{i,j} = 4v_{i,j-1} - v_{i-1,j-1} - v_{i,j-2} - v_{i,j} - v_{i+1,j-1} \quad (65)$$

Равенство нулю правого нижнего элемента этой маски не позволяет в отличие от масок *E* и *K* принимать этот элемент за начало отсчета её элементов *i, j*. В качестве начала здесь выбран правый средний элемент, а расчёт следует вести для элементов в интервалах номеров: $2 \leq i \leq M - 1$ и $3 \leq j \leq N$ (где *M* и *N* – число строк и столбцов изображения, соответственно).

На рисунке 70 приведены изображения: тестовое *v*; масштабированное контурное *L*; контрастированное контурное с обрезкой модуля значений пикселей больше 120 и масштабированное *L*; контурное в псевдоцветах с красными положительными и синими отрицательными значениями контуров *L*. Все изображения на рисунке 70 получены аналогично изображениям на рисунках 68 и 69, но с другой маской и другой верхней границей. Несмотря на различия полученных контурных изображений, все они при известных границах этого изображения могут быть восстановлены до изображения исходного вида. Другими словами, при контурном кодировании изображения следует оставлять без изменения некоторые его границы.

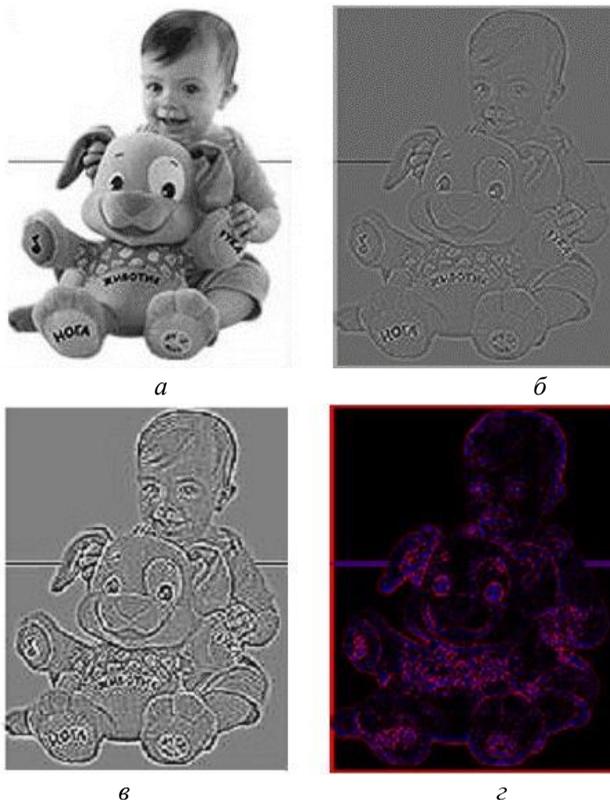


Рисунок 70 – Изображения: исходное $v - a$;
контурное масштабированное $L - б$;

искусственно контрастированное $L - в$; псевдо-цветное контурное $L - г$

В рассмотренных случаях масок размером 3×3 элемента ширина этих границ не превышает двух строк (столбцов), а сами границы не замкнуты. В любом случае границы должны быть такими, чтобы обеспечить вычисление начальных элементов процедуры восстановления.

5.3. Восстановление изображения

Найти алгоритм восстановления (декодирования) исходного изображения $v = Ve$ из контурного E -изображения можно, используя уравнение (61), из которого следует, что:

$$\begin{aligned}
 Ve_{i,j} = & 2Ve_{i-2,j-1} + 2Ve_{i-1,j-2} + 2Ve_{i,j-1} + 2Ve_{i-1,j} + \\
 & + E_{i,j} - 4Ve_{i-1,j-1} - Ve_{i-2,j-2} - Ve_{i-2,j} - Ve_{i,j-2}
 \end{aligned}
 \tag{66}$$

При этом известны граничные значения:

$$\begin{aligned}
 Ve_{1,j} &= E_{1,j} = v_{1,j}, & Ve_{2,j} &= E_{2,j} = v_{2,j}, \\
 Ve_{i,1} &= E_{i,1} = v_{i,1}, & Ve_{i,2} &= E_{i,2} = v_{i,2}
 \end{aligned}$$

и расчёт ведётся для элементов исходного изображения в интервале номеров $3 \leq i \leq M$ и $3 \leq j \leq N$.

Для иллюстрации последовательности декодирования контурного изображения на рисунке 71 *a* показан порядок восстановления: строки выделены тремя различными цветами. При описанном порядке восстановление изображения произойдет за $(M-2) \times (N-2)$ шагов.

Существует и другой порядок восстановления, если имеется возможность параллельной обработки изображения. В этом случае контурное изображение декодируется целой строкой за один шаг обработки. На рисунке 71 *б* такие строки показаны одним цветом и пронумерованы одинаковыми числами. Как видим, эти строки расположены по диагоналям клетчатого поля изображения. При параллельном восстановлении оно произойдет всего за $(M-2)$ шага, если $M > N$, и за $(N-2)$ шага, если $M < N$. Очевиден выигрыш во времени при параллельной обработке контурного изображения, что особенно важно, когда таких изображений много.

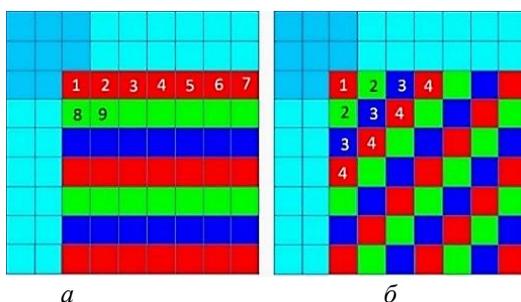


Рисунок 71 – Прямоугольная решетка; восстановление исходного изображения размера 9×9 элементов из контурного того же размера в порядке номеров: последовательное – *a*; параллельное – *б*

Алгоритм восстановления (декодирования) исходного $v = Vk$ изображения из контурного K -изображения можно получить, используя уравнение (62), из которого следует, что:

$$\begin{aligned}
 Vk_{i,j} &= 8Vk_{i-1,j-1} - Vk_{i-2,j-2} - Vk_{i-2,j} - Vk_{i,j-2} - \\
 &- Vk_{i-2,j-1} - Vk_{i-1,j-2} - Vk_{i,j-1} - Vk_{i-1,j} - Vk_{i,j}.
 \end{aligned} \quad (67)$$

При этом известны граничные значения:

$$\begin{aligned}
 Vk_{1,j} &= K_{1,j} = v_{1,j}, & Vk_{2,j} &= K_{2,j} = v_{2,j}, \\
 Vk_{i,1} &= K_{i,1} = v_{i,1}, & Vk_{i,2} &= K_{i,2} = v_{i,2}
 \end{aligned}$$

и расчёт ведётся для элементов исходного изображения в интервале номеров $3 \leq i \leq M$ и $3 \leq j \leq N$.

Таким образом, порядок восстановления из контурного K -изображения отличается от предыдущего случая E -изображения только алгоритмом.

Алгоритм восстановления (декодирования) исходного $v = V_l$ изображения из контурного L -изображения можно получить, используя уравнение (65), из которого следует, что:

$$V_{i,j} = 4V_{i,j-1} - V_{i-1,j-1} - V_{i,j-2} - V_{i+1,j-1} - L_{i,j}. \quad (68)$$

При этом известны граничные значения:

$$\begin{aligned} V_{1,j} = L_{1,j} = v_{1,j}, & \quad V_{M,j} = L_{M,j} = v_{M,j}, \\ V_{i,1} = L_{i,1} = v_{i,1}, & \quad V_{i,2} = L_{i,2} = v_{i,2}, \end{aligned}$$

и расчёт ведётся для элементов исходного изображения в интервале номеров:

$$2 \leq i \leq M-1 \quad \text{и} \quad 3 \leq j \leq N.$$

Заметим, что в алгоритме (68) в отличие от предыдущих алгоритмов (66) и (67) присутствует член с индексом $(i+1)$, что обязывает правильно организовать циклы вычислений. Цикл по индексу i должен быть внутренним. Например, в системе *Mathcad* следует использовать программный модуль:

$$V_l := \left| \begin{array}{l} V_l \leftarrow L \\ \text{for } j \in 3..N \\ \text{for } i \in 2..M-1 \\ \quad V_{i,j} \leftarrow 4 \cdot V_{i,j-1} - V_{i-1,j-1} - V_{i,j-2} - V_{i+1,j-1} - L_{i,j} \\ V_l \end{array} \right. \quad (69)$$

Восстановление исходного изображения из контурного, полученного с применением любой из приведённых на рисунке 67 масок, по приведённым алгоритмам позволяет получить точную копию исходного изображения, приведённого в примерах на рисунках 68 а, 69 а и 70 а.

5.4. Гексагональная решетка

Гексагональные решетки для представления изображений обладают рядом преимуществ по сравнению с квадратными решетками пикселей. В частности, обладая более равномерным соединением пикселей в решетке, они обеспечивают лучшее угловое разрешение благодаря наличию большего числа ближайших

соседей. Они лучше (более гладко) представляют контуры объектов [43], [44]. Ввиду высокой плотности упаковки гексагональные решетки широко распространены в зрительных системах различных организмов. На рисунке 72 а приведена маска из шестиугольных элементов «снежинка», полученная как результат последовательного применения прямой и обратной конечных разностей в трёх различных направлениях осей: i , j , и k , расположенных под углом 60° друг к другу. Такая маска является дискретным аналогом трёхмерной частной производной второго порядка.

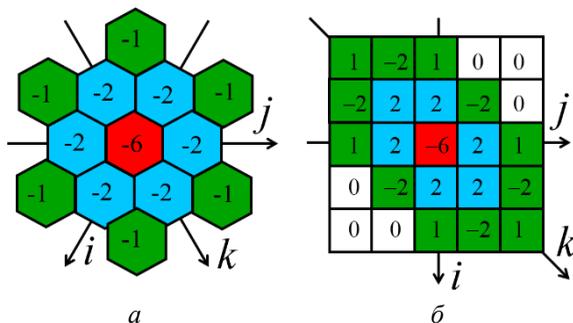


Рисунок 72 – Гексагональная маска дискретного аналога трёхмерной производной второго порядка «снежинка» – а; ортогональная карта «снежинки» – б

На рисунке 72 б приведена прямоугольная карта «снежинки», где её оси сдвинуты до углов 45° . Разумеется, при приведённом плоском картировании неизбежны визуальные искажения, но карта позволяет проводить с ней вычисления и отображать результаты на обычном прямоугольном растре. Другим способом отображения маски рисунка 72 а является её представление в трёхмерном пространстве в виде своеобразного «кубика Рубика», который состоит из 27 меньших кубиков – вокселей. Воксели имеют различные веса: угловые – «1», средние ребра – «-2», средние грани – «4», а центральный – «-8». При этом оси i , j , и k , проходящие через центр «кубика Рубика» и центры боковых граней, взаимно перпендикулярны. Кроме выражений (55), (56) для оси i и (58), (59) для оси j приведём выражения для отображения оси k на карту обратной конечной разности:

$$p_{i,j} = e_{i,j} - e_{i-1,j-1} \tag{70}$$

и прямой конечной разности:

$$s_{i,j} = p_{i+1,j+1} - p_{i,j} \tag{71}$$

Подставляя в последнее выражение значения всех предыдущих, получим аналитическое выражение для двумерной карты гексагональной маски дискретного аналога трёхмерной производной второго порядка рисунка 72 б при $Y = s$, при этом начало отсчёта выберем в правом нижнем углу карты:

$$Y_{i,j} = 2v_{i-3,j-3} + 2v_{i-1,j-1} + 2v_{i-2,j-3} + 2v_{i-2,j-1} + 2v_{i-3,j-2} + 2v_{i-1,j-2} - 2v_{i-3,j-1} - 2v_{i-1,j-3} - 2v_{i-3,j-4} - 2v_{i-1,j} - 2v_{i-4,j-3} - 2v_{i,j-1} - 6v_{i-2,j-2} + v_{i-4,j-4} + v_{i,j} + v_{i-2,j-4} + v_{i-2,j} + v_{i-4,j-2} + v_{i,j-2} \quad (72)$$

На рисунке 73 приведены изображения: тестовое v ; масштабированное контурное Y ; контрастированное контурное с обрезкой модуля значений пикселей больше 120 и масштабированное Y ; контурное в псевдоцветах с красными положительными и синими отрицательными значениями контуров Y .

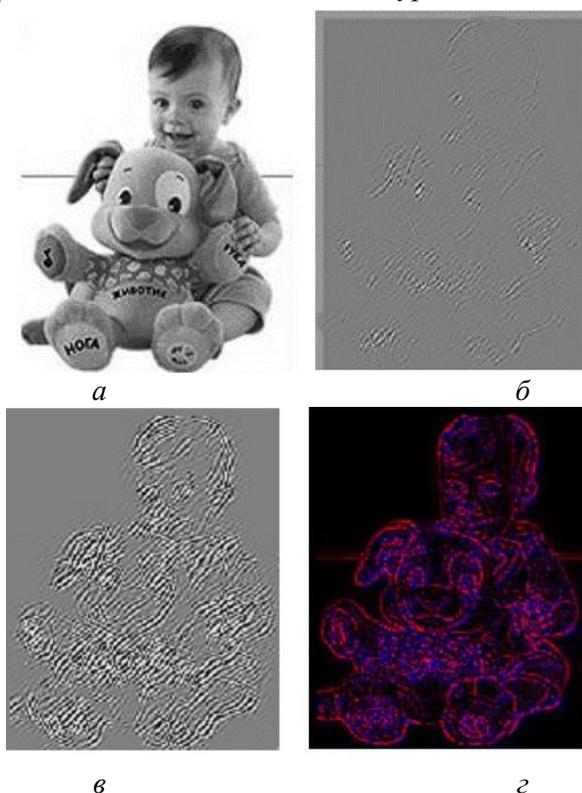


Рисунок 73 – Изображения: исходное v – а; контурное масштабированное Y – б; искусственно контрастированное Y – в; псевдо-цветное контурное Y – г

Все изображения на рисунке 73 получены аналогично изображениям рисунка 68, но с картой гексагональной маски «снежинка». Сравнивая контурные (кодированные) изображения рисунков 68, 69, 70 и 73 можно видеть, что все они различны, несмотря на то что соответствуют одному исходному изображению. Их различие объясняется различием масок, с которыми производится свёртка исходного изображения. Часто для оценки эффективности кодирования изображений используют оценки их информационного сжатия. Например, коэффициент сжатия, представляющий собой отношение размеров выходного и входного файлов кодера или обратная ему величина – фактор сжатия. Эти оценки достаточно трудно сделать, поскольку размер выходного файла зависит от энтропии исходного изображения, которую достаточно сложно вычислить, ввиду её зависимости от вероятностной модели изображения, которую ещё необходимо построить.

С другой стороны, легко оценить интегральную мощность сигналов пикселей исходного и кодированного изображений и далее оценить выигрыш в мощности при кодировании, взяв их отношение. Последняя оценка, хотя и не связана прямо со скоростью передачи информации по каналу связи, позволяет оценить снижение энергетических затрат при такой передаче. В соответствии с вышесказанным выигрыш в мощности определяется выражением:

$$W(I, C) = \frac{\sum_{j=1}^N \sum_{i=1}^M (I_{i,j})^2}{\sum_{j=1}^N \sum_{i=1}^M (C_{i,j})^2}, \quad (73)$$

где I, C – матрицы исходного и кодированного (контурного) изображения; $I_{i,j}, C_{i,j}$ – элементы исходного и кодированного изображений; M, N – число строк и столбцов изображений, соответственно.

Подставляя в выражение (73) матрицы исходного изображения v и масок E, K, L и Y , получим соответствующие выигрыши в мощности:

$$\begin{aligned} W(v, E) &= 9,675; & W(v, K) &= 2,141; \\ W(v, L) &= 8,347; & W(v, Y) &= 2,666. \end{aligned}$$

Согласно этим значениям, наибольший выигрыш достигается при использовании маски E . Что касается маски Y , то расчёт для неё производился по карте маски, а не напрямую и поэтому не подлежит сравнению с остальными масками.

Рассмотрим ещё одну гексагональную маску. На рисунке 74 а приведён её вид, а на рисунке 74 б её прямоугольная карта. Эта маска напоминает графическое изображение рецептивного поля ганглиозной клетки сетчатки глаза, обозначим её буквой R . Алгоритм свёртки исходного изображения v с картой этой маски представляется выражением, которое схоже с выражением (72) за исключением двух угловых элементов, которые в карте равны нулю, и веса центрального элемента:

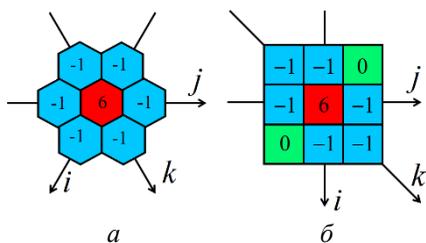


Рисунок 74 – Гексагональная маска «рецептивного поля» R (а); ортогональная карта «рецептивного поля» R (б)

$$R_{i,j} = 6v_{i-1,j-1} + 2v_{i-2,j-2} - v_{i,j} - v_{i-1,j-2} - v_{i-1,j} - v_{i-2,j-1} - v_{i,j-1} \quad (74)$$

Выигрыш в мощности для исходного изображения v и маски R согласно выражению (73) равен: $W(v, R) = 3,685$, что больше, чем для $W(v, Y)$. Заметим, что выигрыш в мощности для биологических систем важен, поскольку даёт им преимущество в выживании.

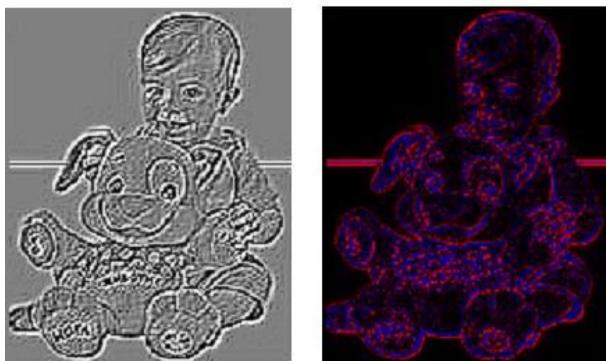
На рисунке 75 приведены изображения: тестовое (v), масштабированное контурное R , контрастированное контурное с обрезкой модуля значений пикселей больше 120 и масштабированное R , контурное в псевдоцветах с красными положительными и синими отрицательными значениями контуров R .



а



б



\bar{b} z
 Рисунок 75 – Изображения: исходное (v) – a ;
 контурное масштабированное R – \bar{b} ;
 искусственно контрастированное R – \bar{b} ;
 псевдо-цветное контурное R – z

Все изображения на рисунке 75 получены аналогично изображениям рисунка 68, но с картой гексагональной маски «рецептивное поле» – R .

5.5. Восстановление исходного изображения

Алгоритм восстановления (декодирования) из Y -контурного изображения исходного изображения $v = Vy$, можно получить, используя уравнение (72), из которого следует, что:

$$\begin{aligned}
 Vy_{i,j} = & Y_{i,j} - 2v_{i-1,j-1} - 2v_{i-1,j-2} - 2v_{i-2,j-1} + 2v_{i-1,j-3} + 6v_{i-2,j-2} + \\
 & + 2v_{i-3,j-1} - 2v_{i-2,j-3} - 2v_{i-3,j-2} - 2v_{i-3,j-3} + 2v_{i-3,j-4} + 2v_{i-4,j-3} + \\
 & + 2v_{i,j-1} + 2v_{i-1,j} - v_{i-2,j-4} - v_{i-4,j-2} - v_{i-4,j-4} - v_{i,j-2} - v_{i-2,j}.
 \end{aligned} \quad (75)$$

При этом известны граничные значения:

$$\begin{aligned}
 Vy_{1,j} = Y_{1,j} = v_{1,j}, \quad Vy_{2,j} = Y_{2,j} = v_{2,j}, \quad Vy_{3,j} = Y_{3,j} = v_{3,j}, \quad Vy_{4,j} = Y_{4,j} = v_{4,j}, \\
 Vy_{i,1} = Y_{i,1} = v_{i,1}, \quad Vy_{i,2} = Y_{i,2} = v_{i,2}, \quad Vy_{i,3} = Y_{i,3} = v_{i,3}, \quad Vy_{i,4} = Y_{i,4} = v_{i,4}
 \end{aligned}$$

и расчёт ведётся для элементов исходного изображения в интервале номеров $5 \leq i \leq M$ и $5 \leq j \leq N$.

Таким образом, порядок восстановления исходного изображения из Y -контурного изображения отличается от предыдущих случаев более широкой границей, что определяется размером карты маски этого изображения.

Алгоритм восстановления (декодирования) исходного изображения $v = Vr$ из R -контурного изображения, можно получить, используя уравнение (74), из которого следует, что:

$$\begin{aligned} Vr_{i,j} = & 6Vr_{i-1,j-1} - Vr_{i-2,j-2} - Vr_{i-2,j-1} - \\ & - Vr_{i-1,j-2} - Vr_{i,j-1} - Vr_{i-1,j} - R_{i,j}. \end{aligned} \quad (76)$$

При этом известны граничные значения:

$$\begin{aligned} Vr_{1,j} = R_{1,j} = v_{1,j}, \quad Vr_{2,j} = R_{2,j} = v_{2,j}, \\ Vr_{i,1} = R_{i,1} = v_{i,1}, \quad Vr_{i,2} = R_{i,2} = v_{i,2} \end{aligned}$$

и расчёт ведётся для элементов исходного изображения в интервале номеров $3 \leq i \leq M$ и $3 \leq j \leq N$.

На рисунке 76 показаны некоторые возможные варианты порядка восстановления небольшого исходного изображения при $M = 10$ и $N = 9$. В варианте на рисунке 76 *a* восстановление начинается с элемента, обозначенного номером «1» и продолжается вдоль третьей строки изображения в порядке следования обозначенных номеров элементов. Эта строка для наглядности выделена тонально (красным цветом). Следующая (четвёртая) строка восстановления, также выделенная тонально (зеленым цветом), начинается с номера «1», продолжается номером «2» и далее до номера «7». Последующие строки восстанавливаются в аналогичном приведённому порядке. При этом восстанавливающая маска движется вдоль строки, касаясь своим нижним элементом очередного элемента изображения. Гексагональная восстанавливающая маска для первого восстановленного элемента изображения выделена тонально и обозначена буквой «Ц» на своем центре. Таким образом производится последовательная (поэлементная) обработка изображения. Верхняя скошенная граница данного варианта изображения имеет толщину в два элемента, боковые в один элемент, а нижняя граница отсутствует.

В варианте на рисунке 76 *б* восстановление также начинается с элемента, обозначенного номером «1», затем «2», затем одновременно восстанавливаются сразу два элемента с номером «3», затем одновременно два элемента с номером «4», затем три элемента с номером «5» и далее последующие горизонтальные, тонально (цветом) выделенные, строки изображения.

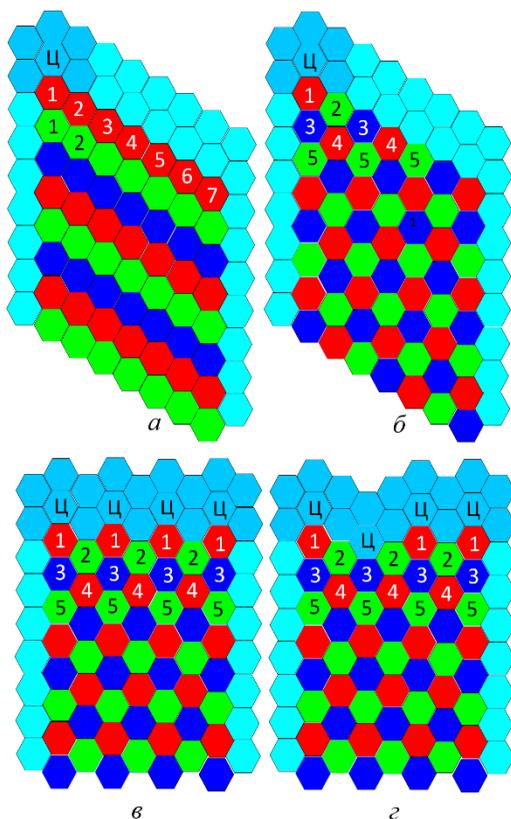


Рисунок 76 – Гексагональная решетка; восстановление исходного изображения размером 10×9 элементов из контурного того же размера: последовательное – *а*; параллельное – *б*, *в*, *г*

Таким образом производится параллельная (построчная) обработка изображения, при движении сверху вниз «волны вычислений», фронт которой совпадает с очередной строкой изображения.

В варианте на рисунке 76 *в* верхняя граница изображения уже не скошенная, а прямая. Обработка изображения ведётся параллельно, построчно, при движении волны вычислений сверху вниз.

В варианте на рисунке 76 *г* верхняя граница изображения не прямая, а со сдвигом вниз одного из её столбцов. Обработка изображения, как и в двух предыдущих вариантах, ведётся параллельно, построчно при движении волны вычислений сверху вниз. Этот вариант показывает, что граница изображения не обязательно должна быть прямолинейной.

Во всех вариантах важно, чтобы каждый из восстанавливаемых элементов изображения в момент восстановления входил элементом в какую-либо маску восстановления так, чтобы значения всех остальных накрываемых маской элементов изображения на момент восстановления были известны.

Отметим выигрыш во времени при параллельной обработке контурного изображения на гексагональной решетке при восстановлении, что особенно важно, когда изображения должны быстро изменяться. Например, для варианта на рисунке 76 в число тактов времени при параллельном восстановлении равно числу восстановленных строк $M = 15$, а при последовательном, при числе элементов в строке $N = 4$, требуется $M \times N = 60$ тактов, т. е. в четыре раза больше, что снижает частоту кадров в четыре раза. При $M \approx N$ выигрыш во времени составляет величину, пропорциональную квадратному корню из числа элементов изображения без границ.

Интересным примером организации решётки изображения представляется зрительная кора головного мозга человека, на которую отображается сетчатка глаза. Она разделена на две половины в левой и правой затылочных областях мозга. Каждая из двух половин зрительной коры при её расположении на плоскости представляет собой слегка усечённый на одном конце эллипс [45], на который проецируется сигнал от одной из двух половин сетчатки, левого или правого поля зрения. На рисунке 77 приведена модель такой коры.

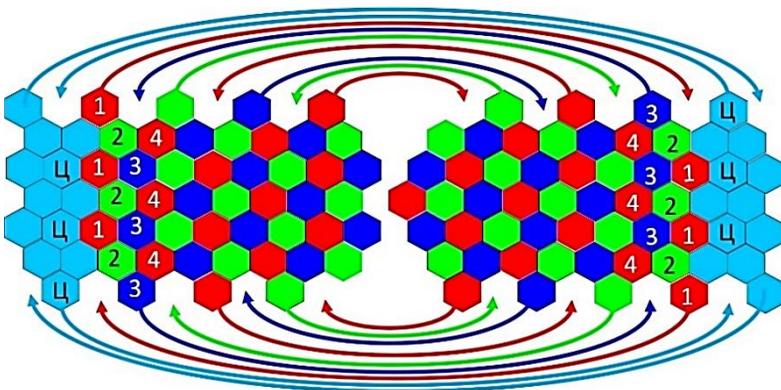


Рисунок 77. Упрощённая модель зрительной коры человека, состоящая из двух половин «склеенных» (в топологическом смысле) в направлениях, указанных стрелками

Усечённые «эллипсы» справа и слева рисунка «склеены» (в топологическом смысле) по направлениям, указанным стрелками так, что элементы в начале стрелок входят в места, указанные их остриями без перекрытия элементов. Аналогом стрелок в мозге являются нервные волокна, соединяющие половины зрительной коры. После склейки образуется поверхность, напоминающая кожуру огурца с обрезанным концом. На эту поверхность топографически проецируется контурное изображение объектов, спроецированных на сетчатку хрусталиком глаза. Преобразование исходного оптического изображения на сетчатке в дискретное контурное осуществляется благодаря центрально-симметричной *on-off* организации рецептивных полей ганглиозных (передающих) нейронов сетчатки [6]. В описанной ситуации мы можем предположить, что контурное изображение может быть восстановлено в исходное методом, рассмотренным в настоящем разделе на ряде примеров. При этом граница изображения представляет собой замкнутое кольцо, соответствующее срезу кожуры «огурца», а кольцеобразный фронт «волны вычислений» распространяется от границы в направлении самой удалённой от границы точки поверхности коры. Заметим, что граница соответствует периферии поля зрения, которая, согласно исследованиям А. Л. Ярбуса [46], освещается равномерно. При таком освещении достаточно передавать на границу зрительной коры только данные о средней освещённости, увеличивая таким образом выигрыш в мощности при контурном кодировании. Наличие волновых процессов в зрительной коре обнаружено И. А. Шевелевым [47].

5.6. Сглаживание локальных дефектов изображения

Изображения могут иметь локальные дефекты, в зоне которых отсутствуют либо элементы изображения (пиксели), либо их сигналы, например, при отсутствии сигналов от соответствующих элементов датчика изображения – сенселей. В системах технического зрения возможны локальные дефекты передающих матриц, а в биологических зрительных системах локальные дефекты сетчатки. Например, в сетчатке глаз человека имеются дефекты, соответствующие местам входа зрительных нервов. Воспринимая изображение окружающей среды, мы не воспринимаем эти дефекты как чёрные пятна, однако можем их обнаружить как «слепые пятна», при наблюдении одним глазом. В этом случае на месте этих пятен ощущается яркость окружающего их фона.

Для устранения заметности дефекта изображения следует интерполировать на его место значение яркости (сигналов) элементов, находящихся на границе дефекта. Возможны различные алгоритмы быстрого сглаживания локальных дефектов. На рисунке 78 приведены изображения: тестовое без дефектов и тестовое с дефектом возле левой брови игрушки; со сглаженным дефектом методом верхней релаксации при итерационном алгоритме сглаживания Vd и со сглаженным дефектом методом перекрестной интерполяции Vdd .

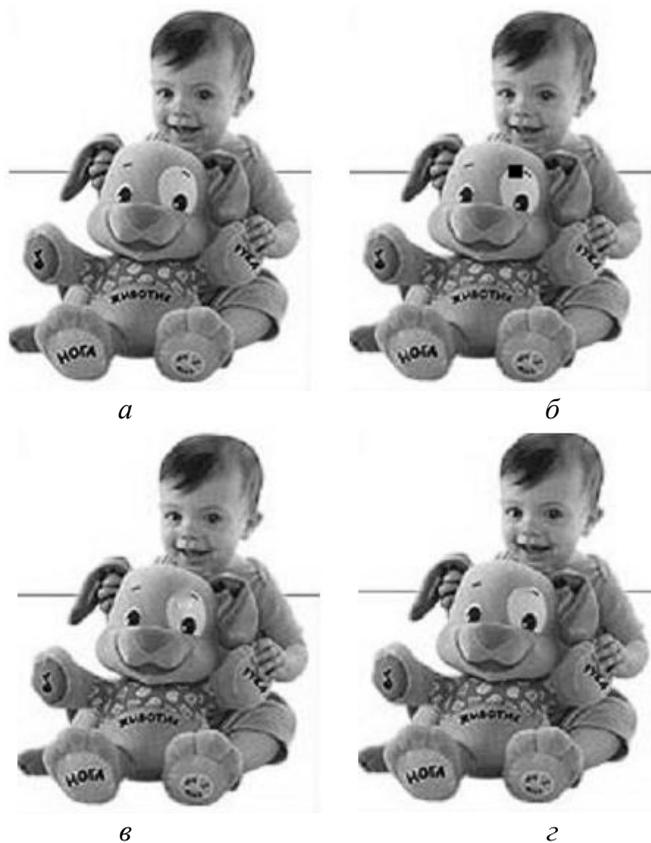


Рисунок 78 – Изображение: исходное без дефектов – а; с дефектом возле левой брови игрушки – б; сглаживание дефекта методом верхней релаксации – в; сглаживание дефекта методом перекрестной интерполяции – г

Итерационный метод верхней релаксации используется, в частности, для численного решения уравнения Пуассона, если известны значения решения на границе (задача Дирихле) [48]. Цифровой аналог такого уравнения приведён выше в выражении (68). Итерационное решение задачи Дирихле методом верхней релаксации для восстановления достаточно большого контурного изображения, как в приводимых здесь примерах, требует сотни итераций (прогонов). Выше был продемонстрирован прямой метод получения исходного изображения за один прогон, а метод итераций при небольшом числе прогонов можно использовать для сглаживания небольшого по площади дефекта изображения.

В системе *Mathcad* для получения изображения Vd со сглаженным прямоугольным дефектом размера $(k + 1) \times (q + 1)$ можно использовать программный модуль:

$$Vd := \left| \begin{array}{l} U \leftarrow v \\ \text{for } z \in 1..Q \\ \quad \text{for } j \in n..n + q \\ \quad \quad \text{for } i \in m..m + k \\ \quad \quad \quad U_{i,j} \leftarrow 0.25 \cdot \omega \cdot (U_{i-1,j} + U_{i,j-1} + U_{i+1,j} + U_{i,j+1}) + (1 - \omega) \cdot U_{i,j} \end{array} \right. , (77)$$

где v – заданное перед модулем изображение с дефектом; Q – число итераций; ω – заданный весовой коэффициент; z – номер итерации; n, m – номер столбца и строки, соответственно, верхнего левого угла дефекта изображения; $m + k, n + q$ – номера строки и столбца, соответственно, нижнего правого угла дефекта; U – переменная в цикле итераций, например, изображение на рисунке 78 в получено за шесть итераций ($Q = 6$) при $\omega = 1,6$.

В итерационном методе сглаживания для вычисления значения каждого элемента сглаженного дефекта используются многие элементы исходного изображения, граничащие с дефектом. В приведённом ниже другом методе «перекрестной» интерполяции значение каждого пикселя сглаженного дефекта интерполируется строго по четырём элементам исходного изображения, граничащим с дефектом и лежащим на вертикали и горизонтали, проходящих через данный пиксель.

Для перекрестной интерполяции с учётом обозначений: $a = n - 1, b = n + q + 1, c = m - 1, d = m + k + 1, h = k + 2, l = q + 2$ и ограничений: $m \leq i \leq m + k, n \leq j \leq n + q$ – интерполянт описывается следующим выражением:

$$\bar{v}_{i,j} = \frac{((b-j)v_{i,a} + (j-a)v_{i,b})(i-c)(d-i) + ((d-i)v_{c,j} + (i-c)v_{d,j})(j-a)(b-j)}{l(i-c)(d-i) + h(j-a)(b-j)} \quad (78)$$

Результат используется при одновременном действии обоих указанных ограничений, а вне дефекта воспроизводится исходное изображение $v_{i,j}$, в результате чего образуется изображение со сглаженным дефектом Vdd , приведённое на рисунке 78 з.

Мы рассмотрели ряд методов контурного кодирования изображений путём свёртки (конволюции) их с различными центрально-симметричными масками на прямоугольной и гексагональной решетках. Все использованные маски обладают нулевым средним значением и все полученные контурные изображения могут быть подвергнуты быстрой деконволюции с восстановлением исходных изображений, если известны их граничные значения (задача Дирихле). При кодировании исходного изображения следует сохранять некоторые его границы, вид которых показан выше для различных сеток и масок изображений. Для других масок можно выбрать сохраняемые границы, руководствуясь изложенным методом выбора. Интересно отметить, что, например, в зрительной системе человека периферия сетчатки глаз засвечивается однородным усредненным излучением, благодаря особой организации их оптической системы [46], то есть уровень граничных значений одинаков и может быть передан в зрительную кору с минимальными энергетическими затратами. Такое задание границ несложно обеспечить и для любых цифровых изображений, искусственно добавляя их к исходному изображению.

Наличие дефектов – локальных зон изображения с отсутствием сигнала изображения – делает актуальной задачу их маскировки (сглаживания). Известны два метода такого сглаживания (интерполяции) значений сигналов изображения в зоне дефекта с использованием значений сигналов от элементов изображения, граничащих с зоной дефекта. Такая задача сглаживания характерна для зрительной системы человека для маскировки так называемых

«слепых пятен» сетчаток глаз. Первый из методов – итерационной интерполяции требует нескольких (шести в примере) циклов (кадров зоны дефекта), но обладает простым алгоритмом цикла итерации. Второй – перекрестной интерполяции обеспечивает сглаживание за один цикл, но обладает более сложным алгоритмом (интерполянт), что влияет на время сглаживания при последовательном вычислении значений элементов области дефекта.

Далее рассмотрим модель оппонентно-цветового кодирования сетчаточного изображения. Это кодирование сочетается с рассмотренным выше контурным кодированием и сильно отличается от цветового кодирования в телевидении. Покажем возможность реконструкции (декодирования) образов объектов наблюдения для зрительной системы путём решения задачи Дирихле с использованием «цветового шаблона». Обсудим возможные механизмы пространственной и цветовой константности зрения, распространения синего цвета на фовеальную область и субдискретизации воспринимаемого изображения объектов окружающей среды.

5.7. Оппонентно-цветовое кодирование изображений

При разработке систем технического зрения в рамках проектов по созданию искусственного интеллекта большое значение имеет бионический подход с ориентацией на зрительные системы различных организмов, созданные в процессе их естественной эволюции. Важным свойством зрения является цветоразличение. Если для таких направлений техники, как цветовоспроизведение и колориметрия может быть достаточным представление о цвете, как чувстве яркости (количественная характеристика) и цветности (качественная характеристика), то для инженерии искусственного интеллекта и физиологии этого явно недостаточно. Здесь необходимо комплексное представление о зрении, как таковом, и цвете, как его свойстве.

Зрение помогает нам создавать в сознании образ (модель) окружающего мира. Зрение – не единственное наше чувство (модальность), а одно из многих, и образ мира в сознании человека многомодален. Каждый из известных органов чувств обладает своими рецепторами (датчиками – преобразователями физических сигналов).

Преобразованные сигналы передаются в мозг для формирования динамичного образа мира. Этот образ помогает субъекту (носителю сознания) реализовать свои цели, планируя будущее на основании собственного опыта. Физическим сигналом воздействия на глаз является свет – небольшой участок электромагнитного спектра волн, характеризуемый яркостью, угловым распределением яркости и формой спектра. Свет в диапазоне длин волн от 380 нм (фиолетовый) до 780 нм (красный) считается видимым. Такие волны занимают частотный диапазон от 400 до 790 терагерц (ТГц). Если изменение яркости по полю зрения обычно не превышает двух порядков при изменении коэффициентов отражения объектов наблюдения, то яркость видимого излучения может изменяться в гигантском диапазоне, приблизительно от 10^{-7} до 10^8 кд/м², как показано на рисунке 25, на котором можно видеть поддиапазоны работы различных фоторецепторов сетчатки глаза: палочек, работающих при низких освещённостях, и колбочек, работающих при высоких освещённостях. И те, и другие способны к адаптации в очень широких диапазонах яркости. Адаптация фоторецепторов обеспечивается в основном биохимическим механизмом фототрансдукции [19]. Скорость адаптации очень высока. Оценить эту скорость можно по скорости исчезновения изображения, предъявляемого глазу при искусственной стабилизации изображения на сетчатке. Уже через 1...2 секунды изображение исчезает даже при стопроцентном контрасте исходного изображения. При этом человек начинает видеть сплошное серое, так называемое, «пустое поле» [8].

Препятствуют исчезновению изображения микродвижения глаз, среди которых: микросаккады, дрейф и тремор. Запись микродвижений глаз приведена на рисунке 21. Высокочастотный тремор, (круговое движение ~ 90 Гц) накладывается на медленные дрейфы (изогнутые траектории). Микросаккады – быстрые (~ 25 мс) скачкообразные движения – обычно возвращают образ к центру визуализации (прямые линии). Диаметр части центральной ямки поля зрения, показанной на рисунке, составляет 0,05 мм (~ 10 угл. мин.).

Другие движения глаз – скачки (саккады) необходимой амплитуды – используются для направления взора на объекты внимания субъекта. Запись таких движений глаз осуществляется с помощью особой техники [8] и приведена на рисунке 7. Для слежения за движущимися объектами используются прослеживающие движения глаз (динамическая фиксация) [49].

Движения глаз имеют определяющее значение не только для перечисленных целей, но и для:

- субдискретизации изображения с целью повышения разрешающей способности зрения;
- распространения синего цвета на область ясного зрения.

Далее рассмотрим достижение этих двух целей с учётом феномена *пространственной константности зрения* [50].

Форма спектра света, отражённого от объекта, зависит как от спектра освещения, так и от спектральных характеристик отражения объекта. На рисунке 79 показан энергетический спектр излучения Солнца в единицах Вт/м²·мкм. На графике видно, что цветовая температура излучения Солнца не совпадает с температурой идеального теплового излучателя – абсолютно чёрного тела с температурой излучения 5900 К, поскольку зависит от фильтрующих свойств атмосферы и положения Солнца на небосклоне.

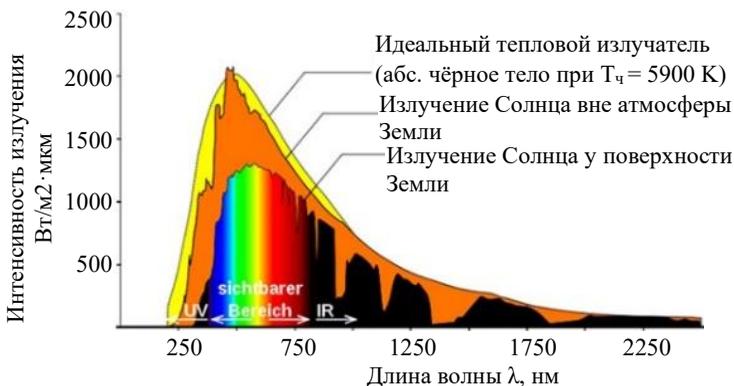


Рисунок 79 – Энергетический спектр излучения Солнца

Существуют различные способы оценки формы спектра отражённого излучения. Например, с помощью спектрозональных телевизионных систем, где количество зон или «цветов» может достигать нескольких сотен. Различать цвета можно по соотношению долей полной световой энергии, прошедших через зональные светофильтры. На рисунке 80 показаны спектральные характеристики отражения контрольных образцов международной комиссии по освещению (МКО).

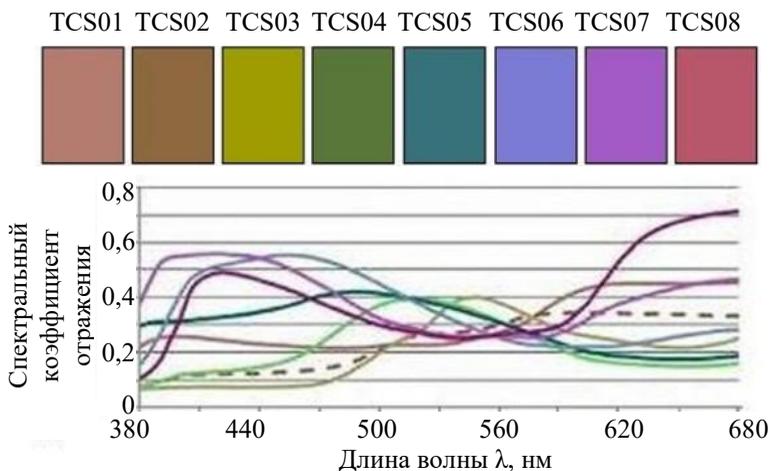


Рисунок 80 – Спектральные коэффициенты отражения 8 контрольных образцов МКО с умеренной насыщенностью цвета

В отличие от зрения животных зрение большинства приматов и человека обладает цветоразличением. У человека обнаружены четыре типа фоторецепторов с различными спектральными характеристиками – относительными спектральными чувствительностями, которые показаны на рисунке 81 сплошными линиями; пятая кривая (пунктир) – так называемая кривая видности глаза определяется суммарными свойствами колбочек.

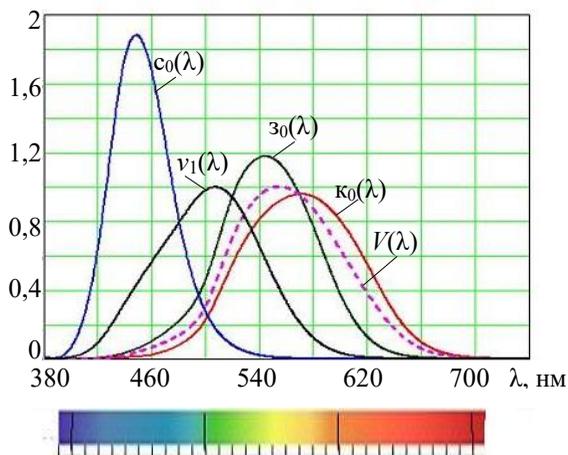


Рисунок 81 – Спектральные характеристики рецепторов сетчатки глаза

Характеристики колбочек: коротковолновая (*S*, синяя), средневолновая (*M*, зеленая), длинноволновая (*L*, красная) перекрывают друг друга. Оставляя в стороне вопрос о работе палочек и взаимодействии палочек и колбочек, отметим только эффект Пуркинье – смещение кривой видности влево до слияния с характеристикой палочек при снижении освещённости. Далее здесь рассматриваем работу колбочек в условиях фотопического зрения.

Абсолютная спектральная чувствительность колбочек не может быть определена из-за упомянутой выше адаптации колбочек в очень большом световом диапазоне. Однако относительная спектральная чувствительность в сравнительно узком диапазоне яркостей может быть определена. Методом уравнивания постоянной энергетической яркости различных по длине волны монохроматических излучений основными излучениями трёхцветного визуального колориметра получают так называемые кривые сложения трихромата (человека с нормальным цветовым зрением), характерные для данного колориметра. Далее, исследуя на этом же колориметре особенности цветовосприятия дихроматов (красно-, зелёно- или, даже очень редких, сине-слепых), по всем полученным данным определяют кривые относительной спектральной чувствительности [51]. Заметим, что в литературе по колориметрии можно видеть различные кривые сложения с отрицательными значениями цвета, которые соответствуют проекции некоторых из основных цветов колориметра на поле измеряемого цвета, а не на поле слагаемых цветов. Форма кривых сложения зависит также от выбранной цветовой модели: *XYZ*, *RGB* и др. Кривые относительной спектральной чувствительности фоторецепторов глаза всегда находятся в области положительных значений чувствительности и не зависят от типа колориметра и модели.

Кривые колбочковых спектральных чувствительностей обычно нормируют по площадям под ними домножением на соответствующие коэффициенты, и тогда при вычислении реакции колбочек на излучение с равномерным энергетическим спектром, эти реакции равны для всех типов колбочек. Одинаковые реакции разнородных колбочек на такой спектр освещения порождают ощущение белого (серого) цвета. Ощущение белого можно также вызвать сочетанием всего трёх первичных цветов колориметра или цветного монитора с характерным для каждого из них отношением яркости этих цветов. Отношения задаются с помощью, так называемых, яркостных коэффициентов цветов.

При изучении зрительной системы мы сталкиваемся с кажущимся противоречием. С одной стороны, *мы видим полноцветную раскраску окружающих нас объектов, а с другой стороны, по зрительному нерву в мозг с сетчатки глаза поступает лишь контурное их изображение*, из-за оппонентной организации зрительных полей ганглиозных нейронов сетчатки [6]. Зрительный нерв содержит более одного миллиона волокон, которые являются аксонами (выходами) ганглиозных нейронов. По каждому волокну передаются сигналы с небольшой, приблизительно круглой, области локализации фоторецепторов сетчатки, причём ганглиозный нейрон суммирует сигналы рецепторов центра круга и вычитает из этой суммы сумму сигналов рецепторов периферии круга. Такая оппонентная структура описывается, например, двумерным оператором Марра ($\nabla^2 G$) [42] с учётом весовой функции суммирования. Вид весовой функции приведён на рисунке 82; для наглядности форма градиционно окрашена снизу вверх, а её ближняя четверть удалена.

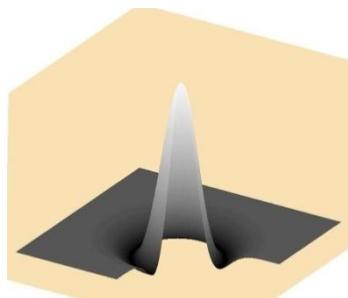


Рисунок 82 – Идеализированная форма весовой функции ганглиозной клетки (нейрона) сетчатки с «*on*» центром и «*off*» периферией

В процессе исследования зрительной системы человека были обнаружены [41] рецептивные поля с различным сочетанием трёх видов колбочек. Их обозначения представлены на рисунке 83. Здесь, кроме обозначений круглых рецептивных полей, приведены обозначения соответствующих масок (крестообразные фигуры) для оппонентно-цветового кодирования тестовых изображений при компьютерном моделировании естественного зрения. Рисунок 83 демонстрирует номенклатуру оппонентно-цветовых рецептивных полей, отличающихся не только сочетанием цветов, но и знаков (+) и (–), которые соответствуют сигналам возбуждения или торможения при активации рецепторов. На представленном рисунке красный цвет означает преимущественное действие в данной зоне рецептивного поля *L*-колбочек, зелёный – *M*-колбочек, синий – *S*-колбочек, желтый – как *L*-, так и *M*-колбочек, белый – всех *L*, *M*, и *S*-колбочек. Рецептивное поле любого нейрона можно определить, измеряя его сигнал при последовательном

освещении рецепторов небольшими пятнами света различного цвета и диаметра круглой или кольцеобразной формы. Такое тонкое исследование с освещением колбочек сетчатки глаза проводилось на обезьянах [6]. Реакции нейронов при различных воздействиях на их рецептивные поля показана на рисунке 84.

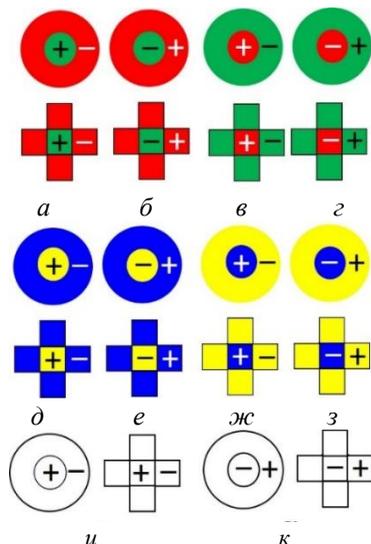


Рисунок 83 – Виды круглых рецептивных полей ганглиозных нейронов сетчатки и их простых крестообразных компьютерных моделей

Рецептивные поля ганглиозных клеток

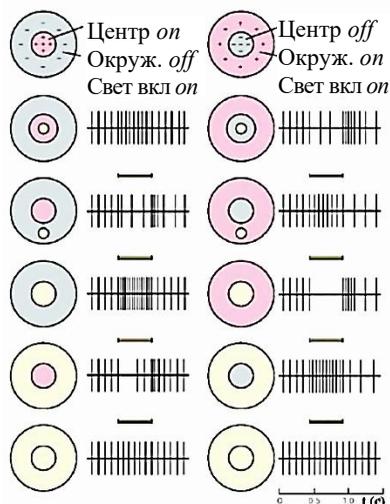


Рисунок 84 – Динамика реакции ганглиозных нейронов на освещение их рецептивных полей пятнами света различной формы

На рисунке 84 слева обозначены реакции нейрона (осциллограммы потенциалов действия – спайков), который возбуждается при включении («on») света в центре его рецептивного поля, а справа реакции нейрона с другим типом рецептивного поля, возбуждающимся при выключении («off») света в центре. Реакции нейронов при аналогичном воздействии на периферию рецептивных полей соответствуют их торможению. Величина реакции пропорциональна частоте спайков справа от изображений рецептивных полей. Положение, размер и форма зоны освещения (обозначена жёлтым цветом) изменяется по строкам изображения. Временной масштаб в правом нижнем углу позволяет сделать вывод о длительностях светового импульса и реакции нейрона. Можно видеть, что длительность реакции порядка одной секунды хорошо

коррелирует со временем появления вышеописанного «пустого поля». Обратное преобразование контурного изображения в изображение полноцветной раскраски возможно при наличии в зрительном тракте механизма восстановления изображения по его контурной версии и известным значениям яркости на его границах, что соответствует решению математической задачи Дирихле [48]. В пользу такого преобразования свидетельствует наличие особого оптического механизма равномерного освещения периферии сетчатки, открытого А. Л. Ярбусом [46].

Особым свойством зрительной системы человека является свойство цветовой константности, когда изменения спектра освещения не влечёт за собой существенного изменения цветности объектов наблюдения, что кажется парадоксальным, учитывая вышесказанное о форме спектра отражённого света.

Далее рассмотрим возможный механизм обеспечения цветовой константности с учётом технологии восстановления контурного изображения. Кажется парадоксальным тот факт, что *в центральной (фовеальной) области сетчатки глаза, соответствующей области ясного зрения, отсутствуют коротковолновые S-колбочки, а присутствуют только средневолновые (M) и длинноволновые (L), и при этом мы ясно видим синие оттенки цвета.*

На рисунке 85 приведена мозаичная модель участка сетчатки глаза [52], которая показывает не только отсутствие синих элементов в центре рисунка, но и наличие довольно больших локальных скоплений однородных по цвету элементов (особенно красных). Далее рассмотрим возможный механизм восприятия цвета с учётом движения глаз, устраняющий кажущийся парадокс.

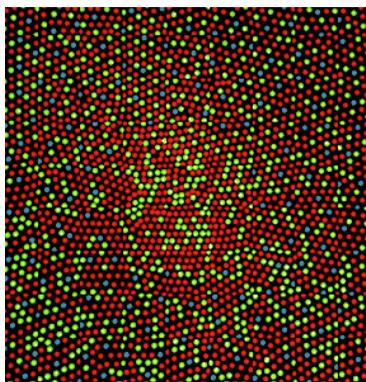


Рисунок 85 – Цветная мозаика распределения разнородных колбочек в фовеальной области сетчатки глаза

Объясняя ранее отмеченные свойства зрительной системы, будем иметь ввиду, что предложенные здесь модели могут быть востребованы при создании систем с искусственным интеллектом, хотя и потребуют разработки соответствующей элементной базы и

необходимого программного обеспечения. В области физиологии зрения эти модели могут приниматься как гипотезы для обоснования проведения соответствующих исследований.

5.8. Дискретизация

Глаз представляет собой многоэлементный приёмник излучения (МПИ). Такие приёмники хорошо известны в технике как матричные фотоприёмники и их даже называют приёмниками «смотрящего типа» [21], в отличие от которых глаз является параллельной системой и не предполагает последовательную съёмку информации с «матрицы». Общим свойством МПИ является пространственная дискретизация изображения, которая, являясь нелинейной операцией, может приводить к появлению в его составе новых пространственных частот. Хорошо известны появляющиеся в этом случае характерные искажения изображения в виде муаров. Появление ложных пространственных частот (алиасинг) особенно заметно при наблюдении периодических текстур с высокой пространственной частотой, сравнимой с частотой дискретизации. Появляющаяся низкочастотная структура вызвана «биением» исходных пространственных частот. Для снижения негативного эффекта дискретизации используются различные методы, учитывающие требования теоремы Котельникова (Найквиста) [11]. К числу таких методов относятся:

- субдискретизация – увеличение числа элементов изображения (пикселей) на приёмной стороне;
- фильтрация – подавление верхних пространственных частот исходного изображения.

В [21] описана процедура «микросканирования» – искусственного приведения в движение оптического изображения на МПИ с помощью вращающихся прозрачных наклонных плоскопараллельных пластин. При этом оптическое изображение не вращается, а только непрерывно сдвигается по круговой орбите. Диаметр орбиты изображения при треморе глаз [8] и микросканировании [21] составляет около половины межэлементного расстояния. Такой сдвиг позволяет удвоить разрешение дискретного изображения по вертикали и горизонтали при его воспроизведении. Для этого следует на воспроизводящей сетке увеличить частоту дискретизации, образуя субпиксели, и коммутировать поочередно субпиксели на включение, синхронно с движением оптического изображения. При этом на прямоугольной сетке в цикле коммутации

задействуются четыре субпикселя, а на гексагональной – три. Гексагональную сетку образуют фоторецепторы сетчатки глаза (см. рисунок 21). На рисунке 86 показан механизм субпикселлизации при гексагональной сетке. Здесь желтое пятно света движется (показаны пять фаз движения) по макросетке фоторецепторов, а на воспроизводящей микросетке три субпикселя, соответствующих одному фоторецептору на сетчатке, переключаются (коммутируются) последовательно на приём сигнала этого фоторецептора синхронно с движением пятна.

На воспроизводящей сетке субпикселей они все объединены в тройки так, что в каждой присутствуют субпиксели первой, второй и третьей фазы включения, раскрашенные на рисунке 86 красным, зелёным и синим цветом. Субпиксели сетки, окрашенные одним цветом (одной фазы), могут включаться одновременно. Фазы чередуются равномерно в каждом цикле.

Если не производить субдискретизацию коммутацией субпикселей, то при микросканировании (треморе) происходит низкочастотная фильтрация с подавлением высших пространственных частот исходного изображения, что помогает устранить алиасинг, но с одновременным снижением чёткости изображения. Заметим, что даже при проведении субдискретизации имеет место некоторое снижение чёткости, обусловленное свёрткой исходного изображения с пространственной функцией чувствительности фоторецептора (сенселя) [11]. В заключение этого раздела необходимо сказать, что при пояснении принципа субдискретизации для упрощения мы абстрагировались от реальной сложности организации зрительной системы и, прежде всего, от свёртки фоторецепторного изображения с весовыми

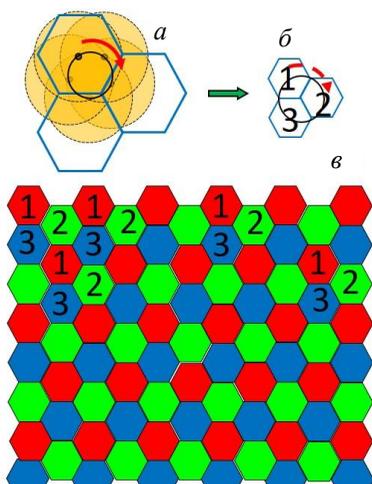


Рисунок 86 – Механизм субпикселлизации при гексагональной сетке: движение пятна света по круговой орбите на сетчатке глаза – *a*; синхронная с движением пятна последовательная коммутация трёх субпикселей (фаз) на воспроизводящей сетке – *б*; примеры троек коммутации с обозначением фазы цветом – *в*

функциями рецептивных полей ганглиозных нейронов сетчатки. В дальнейшем мы вернемся к рассмотрению работы сетчатки глаза, но уже с пониманием содержания термина субдискретизация.

5.9. Пространственная константность

Несмотря на постоянные движения глаз, окружающий мир представляется нам неподвижным (константным). Физиологи зрения иногда предполагают наличие у мозга неких механизмов, компенсирующих движения глаз [50] аналогично компенсации движения головы в приложениях виртуальной реальности [53]. В действительности нет необходимости в компенсации движения глаз для сохранения пространственной константности. При первом взгляде на окружающую среду мозг человека *фиксирует* пространственное положение какого-либо увиденного неподвижного объекта и тем самым задает *начало координат* среды. Последующие смещения глаз или головы человека относительно его тела являются программно-заданными его мозгом, и мозг знает величину этих смещений относительно начала координат. Движения глаз и головы нужны человеку (его мозгу) для нанесения текущего видимого изображения среды на неподвижную «сетку» *её координат с фиксированным* началом. После очередного моргания начало координат может уточняться и даже переноситься на другой неподвижный объект среды. Особым случаем является ситуация, когда положение тела человека меняется произвольным, не зависящим от него способом, например, при его движении в транспорте. В этом случае мозг получает дополнительную информацию о смещениях глаз и головы от вестибулярного аппарата человека и таким образом сохраняет константность мира, обеспечивая неподвижность сетки координат, накладываемой на окружающую среду.

Особенностью зрения человека является его неравномерная разрешающая способность по полю зрения. На рисунке 87 приведена двумерная диаграмма разрешения одного (левого) глаза, из которой видно, что относительное разрешение при удалении от центра фoveальной области сетчатки до $40^\circ - 60^\circ$ постепенно уменьшается в сорок раз.

При такой организации зрения, погрешности коррекции местоположения константной сетки после поворотов, возрастающие при увеличении углов изменения оптической оси глаза незаметны для зрителя. Моделью нанесения текущего видимого изображения

среды на неподвижную «сетку» может служить процесс динамического фовеального рендеринга (DFR), показанный на рисунке 55. Для полноты впечатления следует использовать анимацию gif- файла такого рисунка [36].

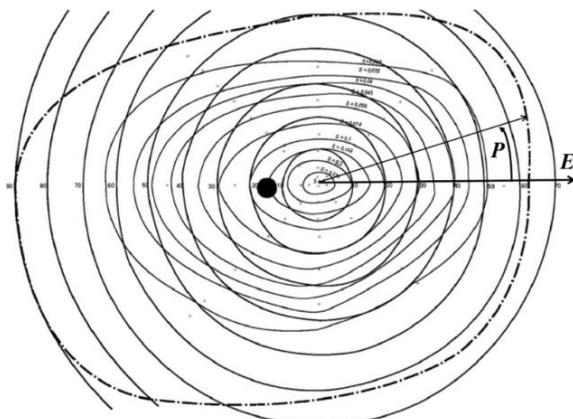


Рисунок 87 – Двумерная диаграмма разрешения глаза

В связи с большой неравномерностью разрешающей способности по полю зрения нанесение текущего видимого изображения среды на неподвижную «сетку» является нерегулярным процессом, тесно связанным с важным для точного распознавания рассматриванием деталей изображения посредством проекции их в фовеальную (ясного зрения) область сетчатки глаза (см. рисунок 7). Область бокового зрения с малым разрешением важна для предварительного распознавания и своевременного обнаружения движущихся объектов.

5.10. Распространения синего цвета на область ясного зрения

В процессе движения глаз одной и той же точке пространства соответствуют различные точки сетчатки глаза. Одной и той же неподвижной детали объекта соответствует различное разрешение глаза в различные моменты времени. Деталь воспроизводится в сознании с некоторым разрешением, зависящим от числа суммированных последовательных «кадров» (моментов) зрения и близости изображения детали к фовеоле (центральной ямке сетчатки глаза). Для периодического тест-объекта (рисунок 88 а) снижение разрешения приводит к уменьшению амплитуды воспринимаемого профиля яркости объекта, синусоидальный график которого показан на рисунке 88 б.

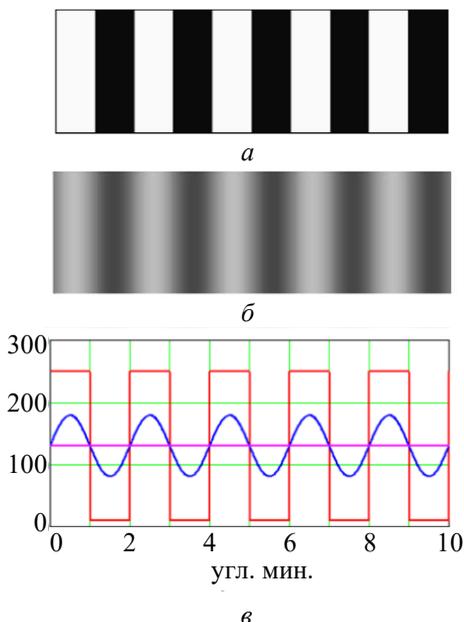


Рисунок 88 – Тест-объект (а), его восприятие (б), соответствующие графики их профилей (в)

Согласно рисунку 84 коротковолновые S-колбочки находятся на расстоянии порядка 10 угловых минут от центра фовеолы, что соответствует длине микросаккады (см. рисунок 21). На таких расстояниях разрешение глаза изменяется незначительно, и с высокой долей вероятности, благодаря частым микросаккадам и дрейфу глаз, в любой точке пространства в зоне ясного зрения появляются рецептивные поля с коротковолновыми «синими» колбочками. Заметим, что разрешение по синему цвету всё же меньше, чем по красному или зелёному благодаря тому, что рецептивные поля, содержащие коротковолновые колбочки несколько больше по размеру, что в свою очередь соответствует меньшему количеству этих колбочек. Описанный механизм распространения синего цвета на область ясного зрения соответствует подобному механизму ночного (скотопического) зрения. Отсутствие палочек в фовеальной области сетчатки практически не сказывается на чувствительности зрения, т. е. качество восприятия сохраняется и при плохом освещении.

5.11. Контурное изображение объектов наблюдения

Сетчатка глаза имеет сложное многослойное строение. Рецепторы через биполярные и другие нейроны (клетки) связаны с ганглиозными нейронами. Горизонтальные связи между клетками сетчатки обеспечивают горизонтальные и амакриновые клетки. В результате взаимодействия клеток сетчатки каждая из ганглиозных клеток имеет связь с множеством рецепторов, образуя своё рецептивное поле (РП). Суммирование сигналов фоторецепторов производится с различными коэффициентами (весами) в зависимости от удалённости фоторецептора от центра РП. По форме РП близко к кругу; в центре круга веса наибольшие, затем, по мере удаления от центра РП, они уменьшаются, достигая нуля на некотором удалении, а затем увеличиваются, имея противоположный знак по сравнению с весами центра и, наконец, снова уменьшаются по абсолютной величине, обращаясь в нуль на краю РП. Рецептивные поля сильно перекрываются между собой [40], как показано на рисунке 89.

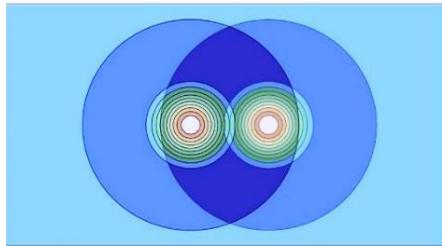


Рисунок 89 – Перекрывание рецептивных полей ганглиозных нейронов; вид сверху на весовые функции в «географической» раскраске

Веса рецепторов РП можно считать выборками двумерной весовой функции. На рисунке 82 показан вид такой функции с положительным центром и отрицательной периферией относительно нулевой плоскости, с вырезанным для наглядности квадрантом. Такой «экзотический» вид весовой функции РП не случаен. Он позволяет оптимальным образом минимизировать искажения, связанные с дискретным характером изображения, передаваемого глазом в зрительную кору головного мозга. В [42] предложено выражение для весовой функции РП, показанной на рисунке 82, которое известно, как оператор (фильтр) Марра:

$$\nabla^2 G(x, y) = \frac{-1}{\pi\sigma^4} \left(1 - \frac{x^2 + y^2}{2\sigma^2} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}}, \quad (79)$$

где x, y – локальные угловые координаты, с началом в центре РП, σ – параметр, определяющий ширину гауссовой функции;

$$G(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}; \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \text{ – оператор Лапласа.}$$

Это выражение удобно тем, что позволяет задать размер РП с помощью параметра σ . Воздействие оператора Марра на оптическое сетчаточное изображение $V(x, y)$ описывается с помощью операции свёртки (конволюции):

$$E(x, y) = \nabla^2 G(x, y) * V(x, y) = \nabla^2 (G * V), \quad (80)$$

Последний член выражения напоминает об ассоциативности свёртки. Выражение (80) описывает аналоговое изображение, которое следует дискретизировать, сделав выборки его значений в заданных точках. Дискретизацию можно представить переходом от координат x и y к индексам i и j в соответствии с выражением:

$$E_{i,j} = E(x, y) \sum_i \sum_j \delta(x - x_i) \delta(y - y_i), \quad (81)$$

где $\delta(x - x_i) = 1$, если $x = x_i$, 0 – иначе. Заметим, что описанная процедура представления сигнального изображения на уровне ганглиозных нейронов не кажется очевидной, если не учитывать ассоциативность свёртки и минимизацию длины связей этих нейронов. С учётом последнего, можно принять расположение точек выборки в местах расположения ганглиозных нейронов. В общем случае сетку дискретизации можно представлять различной конфигурацией, в частности гексагональной и не обязательно равномерной или прямолинейной. Однако моделировать основные процессы преобразования изображений в зрительной системе удобно на прямоугольной сетке, используя методы обработки цифровых изображений.

Аналогом операции (80) при обработке цифрового изображения является его свёртка с некоторой маской (фильтром, ядром, керном, матрицей, оператором, весовой функцией). На рисунке 90 приведён пример масок размером 3×3 элемента. Здесь цветом выделены различные значения весов ячеек маски. Сумма весов каждой маски равна нулю. Любая из таких масок может использоваться для моделирования рецептивного поля. Заметим, что на рисунке 90 изображены маски с положительными весами в центрах, которые соответствуют РП с «он» центром.

-1/4	-1/2	-1/4
-1/2	1	-1/2
-1/4	-1/2	-1/4

-1/8	-1/8	-1/8
-1/8	1	-1/8
-1/8	-1/8	-1/8

0	-1/4	0
-1/4	1	-1/4
0	-1/4	0

a
б
в

Рисунок 90 – Маски фильтров: оператор дискретного аналога двумерной производной второго порядка – *a*; однородно-периферийный оператор – *б*; дискретный аналог оператора Лапласа – *в*

Для моделирования РП с «*off*» центром все веса масок следует инвертировать (умножить на минус единицу).

Особо отметим то, что из-за различных размеров РП ганглиозных нейронов по площади сетчатки, сетка дискретизации сильно разрежена по мере удаления от центральной ямки. Однако, при отображении на зрительную кору, например (*V1*), она сильно деформируется так, что приобретает равномерность. С учётом эффекта разреженности пространственной константности мы воспринимаем исходное изображение рисунка 91 *a* как показанное на рисунке 91 *б*, где чётко виден лишь левый глаз игрушки, а большая часть изображения размыта тем больше, чем дальше она от области ясного зрения. То есть нанесение изображения на пространственную сетку, как говорилось выше, снимает деформацию, но оставляет размытие. Рисунок 91 *в* даёт представление о виде деформации изображения на первичной зрительной коре, хотя и не соответствует реальному коду изображения. Далее моделирование свёртки весовой функции рецептивного поля и изображения можно вести без учёта деформации и размытия изображения в зрительном тракте, которые нетрудно учесть, используя методы обработки изображений, применённые при создании рисунка 91.



Рисунок 91 – Исходное изображение – *a*; наблюдаемое зрительной системой – *б*; иллюстрация деформации изображения на первичной зрительной коре – *в*

При моделировании свёртки весовой функции рецептивного поля и изображения будем применять маску – дискретный аналог оператора Лапласа, показанную на рисунке 90 в. Использование других масок описано в [54] и выше.

Математическое выражение для алгоритма свёртки исходного изображения с маской, представленной на рисунке 90 в, содержит всего 5 членов, поскольку угловые коэффициенты этой маски равны нулю, и имеет вид:

$$L_{i,j} = V_{i,j-1} - 0,25(V_{i-1,j-1} + V_{i,j-2} + V_{i,j} + V_{i+1,j-1}). \quad (82)$$

Обратим внимание на то, что в выражении (82) член $V_{i,j}$ соответствует правому среднему элементу маски рисунка 90 в, но не центру маски, как можно было бы ожидать из соображений симметрии. Такая индексация удобна для операции последующего восстановления свёрнутого изображения до исходного вида. Операция восстановления называется *деконволюцией*. На рисунке 92 приведено исходное чёрно-белое изображение и вид его свёртки с маской преобразования Лапласа, выполненной в соответствии с выражением (82). Заметим, что в данном случае свёртка приводит к образованию как положительных, так и отрицательных значений элементов $L_{i,j}$, а поскольку элементы картинки должны быть только положительными, результат свёртки следует подвергнуть

процедуре масштабирования с переносом в диапазон яркостей пикселей от 0 до 255.

Изображение рисунке 92 б имеет «контурный вид», который характерен для использования масок рисунка 90 [54]. На рисунке 92 в показана модель деформированного контурного изображения на первичной зрительной коре в условиях скопического (чёрно-белого) зрения. Очевидно, что такое контурное изображение не соответствует воспринимаемому

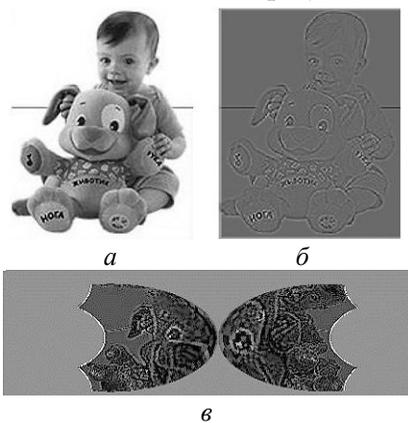


Рисунок 92 – Исходное изображение – а; контурное изображение – б; иллюстрация деформации изображения на первичной зрительной коре – в

в сумерках образу даже с учётом топологических искажений (воспринимаемый образ неконтурный).

Напомним, что нумерацию элементов цифрового изображения принято вести из его левого верхнего угла. Здесь номера строк: $1 \leq i \leq M$, а номера столбцов: $1 \leq j \leq N$, хотя часто нумерацию начинают с нуля. Свёртка в соответствии с выражением (82) проводится начиная с элемента с индексами: $i = 2, j = 3$ и заканчивая элементом с индексами: $i = M - 1, j = N$ в двух циклах: внутреннем по i и внешнем по j . При этом элементы с индексами: $i = 1, M$ и $1 \leq j \leq N$, а также $j = 1, 2$ и $2 \leq i \leq M - 1$ остаются неизменными. Эти элементы составляют *границу* изображения.

Знание границы позволяет осуществить деконволюцию свёрнутого изображения в соответствии с выражением:

$$V_{i,j} = 4V_{i,j-1} - V_{i-1,j-1} - V_{i,j-2} - V_{i+1,j-1} - 4L_{i,j}, \quad (83)$$

которое прямо следует из уравнения (82). Процесс деконволюции осуществляется в том же порядке, как и процесс конволюции (свёртки) в двух циклах: внутреннем по i и внешнем по j с теми же началом и концом. Программный модуль деконволюции (*Machcad*) аналогичен выражению (69).

Если процесс деконволюции всегда последовательный (хотя может быть и распараллелен до осуществления в одном цикле), то конволюция может быть выполнена одновременно во всех элементах изображения, кроме границы. Такая одновременная конволюция (свёртка) исходного изображения с весовыми функциями РП ганглиозных клеток производится на сетчатке глаза. Для объяснения естественного (не контурного, интегрального) восприятия образа окружающего мира мы должны предположить *наличие в мозге механизма деконволюции с непрерывным следованием волн вычислений этого образа из его контурного прообраза*. Волны с подходящими характеристиками в зрительной коре обнаружены [47] без объяснения причины их наличия.

Заметим, что контурный образ удобен для распознавания объекта наблюдения. Можно предположить, что задачи наблюдения и распознавания мозгом объектов разделены и осуществляются параллельно.

В общем случае прямоугольной маски с произвольными весовыми коэффициентами $a_{i,j,k,q}$ свёрнутое изображение определяется выражением:

$$E_{i,j} = a_{i,j,0,0}V_{i,j} + \sum_{k=1}^K a_{i,j,k,Q}V_{i-k,j} + \sum_{q=1}^Q a_{i,j,K,q}V_{i,j-q} + \sum_{q=1}^Q \sum_{k=1}^K a_{i-k,j-q,k,q}V_{i-k,j-q} \quad (84)$$

где $0 \leq k \leq K$, $0 \leq q \leq Q$ – номера ячеек матрицы, отсчитываемые от её правого нижнего угла.

Из выражения (84) следует, что деконволюцию следует осуществлять в соответствии с выражением:

$$V_{i,j} = \frac{1}{a_{i,j,0,0}} \left\{ E_{i,j} - \sum_{k=1}^K a_{i,j,k,Q}V_{i-k,j} - \sum_{q=1}^Q a_{i,j,K,q}V_{i,j-q} - \sum_{q=1}^Q \sum_{k=1}^K a_{i-k,j-q,k,q}V_{i-k,j-q} \right\}, \quad (85)$$

причём, если деконволюцию проводить из левого верхнего угла изображения, то первым восстановленным элементом исходного изображения будет пиксель с индексами $i = K+1, j = Q+1$. При этом должна быть известна верхняя граница изображения с индексами $1 \leq i \leq K$ и левая с индексами $1 \leq j \leq Q$. Аналогично можно определить такой пиксель при начале из любого угла изображения, как внутренний угловой элемент маски и соответствующие границы изображения, перекрывающие угловую маску за исключением данного пикселя. Из выражения (85) следует, что при моделировании РП прямоугольной маской коэффициент $a_{i,j,0,0}$ следует выбирать отличным от нуля.

Перейдём теперь от обсуждения общих вопросов контурного кодирования и декодирования к контурному кодированию цвета. В технике кодирование цвета начинается с матричного фотоприёмника, где отдельные его фотоэлементы – сенсели покрыты оптическими светофильтрами различных цветов. Распределение светофильтров по поверхности матрицы может быть различным у фотодатчиков разных производителей. На рисунке 93 *a* приведён фильтр (шаблон) Б. Э. Байера (сотрудника фирмы Кодак). Этот фильтр содержит элементные светофильтры трёх цветов: красного, зеленого и синего. Фильтр Байера представляет собой регулярную структуру, в которой на два зелёных светофильтра приходится по одному красному и синему. Из электрических сигналов сенселей матричного фотоприёмника формируется видеосигнал, передаваемый на воспроизводящее изображение устройство – монитор. Для лучшей детализации воспроизводимого цветного изображения следует распространить значения сигналов сенселей с красными

светофильтрами на местоположения ближайших к ним зелёных и синих светофильтров и аналогично – значения сигналов «зелёных» сенселей на местоположения «красных» и «синих», а «синих» на местоположения «красных» и «зелёных». На рисунке 93 кроме фрагмента фильтра Байера показаны варианты интерполяции значений цветных сигналов сенселей на соседние местоположения, фрагмент трёхкомпонентного сигнала изображения в результате интерполяции, а также формирование сигналов субпикселей изображения монитора (слева) по сигналам пикселей трёхкомпонентного изображения и фрагмент изображения монитора (справа). Между стадиями «в» и «г» рисунка 93 выполняется передача трёхкомпонентного сигнала по линии связи видекамера (передатчик) – монитор (приёмник). На этой стадии трёхкомпонентный сигнал подвергается кодированию в начале линии и декодированию в конце с различными целями, среди которых: сжатие информации, снижение энергетических затрат передачи, уменьшение возможных искажений сигналов и др. Преобразование сигналов матричного фотоприёмника в трёхкомпонентный производится в видеокамере с последующим универсальным кодированием для любых шаблонов светофильтров.

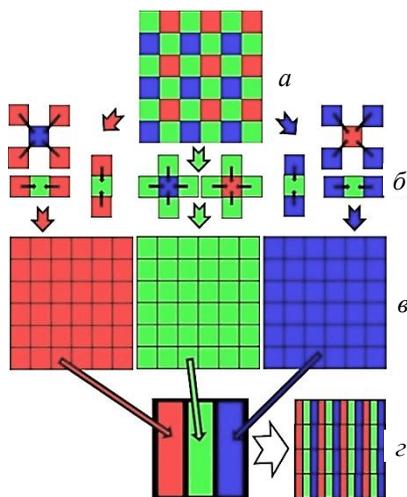


Рисунок 93 – Формирование изображения с использованием шаблона Байера: шаблон – а; интерполяция сигналов цвета – б; фрагмент трёхкомпонентного сигнала изображения – в; субпиксели и фрагмент экрана монитора – г

В зрительной системе функцию сенселя с цветным светофильтром выполняет колбочка. Интерполяция сигналов колбочек с целью получить цветные компоненты высокого разрешения не осуществляется. Вместо этого формируются сигналы рецептивных полей, организованных по принципу оппонентности (противоположности). Например, сигналу центральной L-колбочки противопоставляется (вычитается из него) сумма сигналов,

окружающих M -колбочек (R^+G^- -оппонентность, «он» центр R , «off» окружение G). В обозначении первая буква определяет «цвет» центра РП, вторая – периферии, «+» соответствует «он», а «-» – «off» реакции. Другой пример: сигналу центральной S -колбочки противопоставляется сумма окружающих M - и L -колбочек (B^+Y^- оппонентность). Такое формирование называем *оппонентно-цветовым кодированием* изображения. Если центр рецептивного поля и его периферия формируются из равно взвешенных сигналов всех видов колбочек (L , M и S), то такую оппонентность называем чёрно-белой (W^+K^-) или ($W-K^+$), имея ввиду, что белый и чёрный обозначают разновидности реакции на белую (серую) цветность объектов наблюдения. Обозначения РП этих и других примеров приведены на рисунке 83.

Сигналы собственных рецептивных полей формируют ганглиозные нейроны сетчатки в виде коротких импульсов – потенциалов действия (спайков) на своих аксонах, составляющих волокна зрительного нерва. Величина сигнала кодируется частотой спайков, что в технике называется частотно-импульсной модуляцией. Совокупность сигналов ганглиозных нейронов представляет собой *сигнальное изображение*. Передача сигнальных изображений глаз в кору головного мозга осуществляется через промежуточные звенья зрительного тракта. Промежуточные сигнальные изображения называются проекциями на соответствующие части мозга. Известны проекции на наружные колленчатые тела (НКТ), ядра четырёххолмия, зрительную кору и др. Зрение участвует в организации различных функций мозга: отображения окружающей среды, распознавания объектов среды, предсказание и планирование событий в среде, управления исполнительными органами организма. Промежуточные и корковые проекции сигнальных изображений обладают топографической организацией (ретинотопией) и располагаются в местах, обеспечивающих минимизацию длины нейронных связей.

Выберем в качестве простой модели сетчаточного изображения шаблон Байера, предполагая, что красный светофильтр R соответствует красночувствительному центру РП ганглиозного нейрона и зеленочувствительному G крестообразному окружению, то есть моделируется R^+G^- оппонентность. Синий светофильтр B соответствует синечувствительному центру РП и зелёночувствительному (вместо жёлточувствительного Y) крестообразному окружению, то есть моделируется B^+Y^- оппонентность.

Зелёный светофильтр соответствует бело-чувствительному W центру и периферии РП. Некоторое несоответствие приведённым на рисунке 83 видам РП с точки зрения иллюстрации процесса кодирования не принципиально.

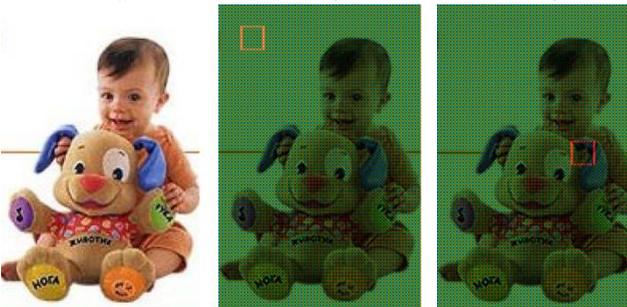
На рисунке 94 *а, б, в* приведены три цветные компоненты $R, G,$ и B цветного изображения (z).



а

б

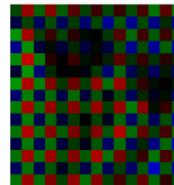
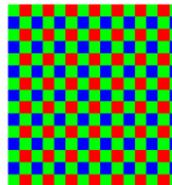
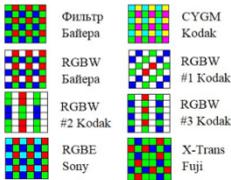
в



з

д

е



ж

з

и

Рисунок 94 – Формирование модели сетчаточного изображения с использованием шаблона Байера: трёхкомпонентное изображение – *а, б, в*; видимое изображение – *г*; «байеровская» модель сетчаточного изображения – *д*; её фрагмент в рамке – *з*; то же самое с другим положением рамки – *е, и*; фрагменты других шаблонов – *ж*

Выборкой элементов этих компонент составляются две одинаковых матрицы Байера (d) и (e) с элементами, подкрашенными в цвет соответствующих компонент. Рисунок 94 з показывает с увеличением участок изображения (d), обведённый рамкой. Рисунок 94 и показывает с увеличением участок в рамке изображения (e). Рисунки ($з$), ($и$) показывают и шаблон Байера (см. рисунок 93 а), и модуляцию яркости пикселей изображения. Заметим, что *шаблон имеет постоянную конфигурацию и не зависит от содержания (модуляции) пикселей изображения*. На рисунке 94 ж приведены конфигурации шаблонов, отличные от шаблона Байера, использованные в разработках датчиков изображений различных фирм. Любую из этих конфигураций, как и любую цветовую конфигурацию сетчатки, можно смоделировать используя цветовые компоненты выбранного изображения и вышеописанный метод. При этом для моделирования периодической конфигурации удобно использовать условные операторы, а для непериодической (как на рисунке 85) можно воспользоваться таблицей размером с изображение.

Сигнал элемента изображения v шаблона Байера может быть формализован с использованием условных операторов (*if* – если) в обозначениях *Mathcad* согласно выражению:

$$v_{i,j} := \text{if} \left[(-1)^{i+j} = 1, G_{i,j}, \text{if} \left[(-1)^i = -1 \vee (-1)^j = 1, R_{i,j}, B_{i,j} \right] \right] \quad (86)$$

Сигналы колбочек глаза и сенселей телекамеры не содержат информации об их спектральных характеристиках, при этом *местоположение ганглиозного нейрона в сигнальном изображении таких нейронов содержит информацию о структуре его РП*.

Сигнальное изображение, полученное методом оппонентно-цветового кодирования исходного изображения, может быть описано свёрткой исходного изображения с весовыми функциями рецептивных полей при учёте цветовой структуры этих полей и последующей выборкой значений свёртки в узлах сетки дискретизации. Учёт цветовой структуры означает наличие информации о типе и весовых коэффициентах каждого рецептивного поля не только на передающем, но и на приёмном конце линии связи (зрительного нерва). Необходимость учёта цветовой структуры приводит к необходимости настройки линии связи в процессе обучения зрению при развитии организма. Первичная (грубая) настройка местоположения концов зрительного нерва происходит при внутриутробном развитии с помощью градиентных молеку-

лярных механизмов в процессе роста нерва, а вторичная (тонкая) – в младенческом возрасте в процессе обучения нейронной сети проекционного экрана первичной зрительной коры на основе принципа Хебба [55]. Обучение сети экрана возможно, поскольку в начале обучения каждый элемент (нейрон) экрана связан синапсами своих дендритов с аксонами всех ближайших волокон зрительного нерва. В процессе обучения синаптические контакты с топографически близкими (на сетчатке) волокнами усиливаются, а с далекими обрываются. Контурный характер сигнального изображения зрительного нерва способствует формированию одновременных сигналов у близких волокон в процессе движения глаз при обучении. Одновременность – основа принципа обучения нейрона по Хеббу, который является аналогом принципа условного рефлекса И. П. Павлова. При настройке экрана возникает чёткая топографическая структура переданного изображения и, в частности, возможность деконволюции этого изображения с восстановлением исходного. На восстановленное изображение может быть «наложен» шаблон исходного, поскольку он не изменяется, и таким образом получено соответствие между спектральной характеристикой центра РП ганглиозного нейрона и видом нейрона с проекционного экрана восстановленного изображения. Возможных видов этих нейронов столько, сколько видов рецептивных полей. Нейроны каждого вида составляют свое дискретное изображение – цветовую компоненту восстановленного изображения, которая может быть выделена или нет в отдельный слой. Возможны любые преобразования компонент, в частности, их субпикселизация интерполяцией, аналогично показанной на рисунке 93 б. Субпикселизацию с интерполяцией можно провести с использованием вышеупомянутого принципа Хебба.

На рисунке 95 приведена модель эволюции сетчаточного изображения в последовательности рисунков от «а» до «е».

Заметим, что рисунок с «наложением шаблона» – это лишь способ подчеркнуть наличие информации о местоположении различных центров РП или соответствующих «пикселей» в сигнальном изображении.

Математически свёртка изображения с оператором Лапласа описывается выражением (82) при подстановке в него вместо элементов $V_{i,j}$ элементов $v_{i,j}$, определяемых выражением (86).

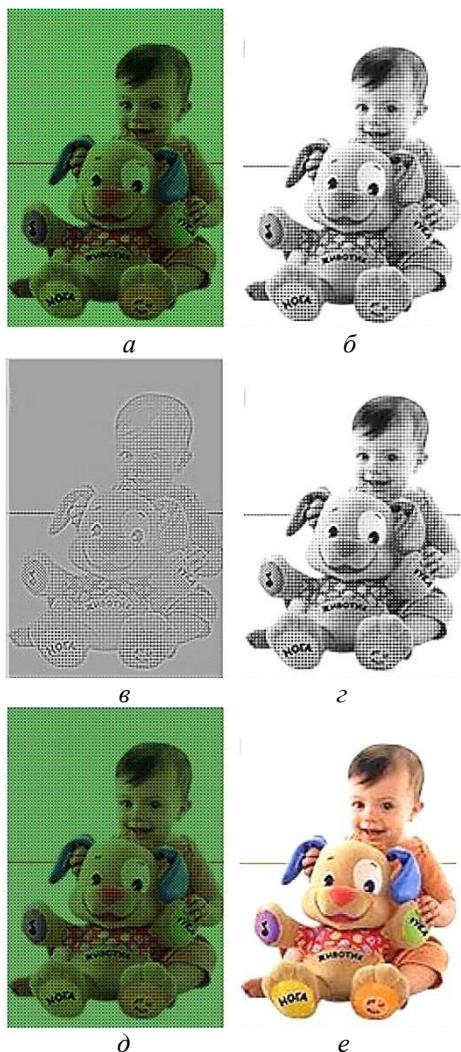


Рисунок 95 – Модель эволюции сетчаточного изображения: сигнальное изображение с наложенным шаблоном светофильтров – *a*; сигнальное изображение – *б*; свёртка (конволюция) сигнального изображения с весовой функцией РП – *в*; деконволюция свернутого изображения эквивалентная изображению «*б*» – *г*; деконволюция с наложенным шаблоном – *д*; воспринимаемое изображение после интерполяции – *е*

Деконволюция (восстановление) свернутого изображения L проводится получением изображения $Vl = v$ с использованием уравнение (83). При этом известны граничные значения:

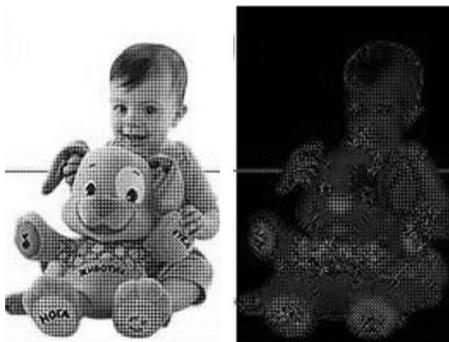


Рисунок 96 – Уменьшение средней мощности (яркости) исходного сигнального изображения (а) при его свёртке с маской Лапласа; модуль свёртки (б) выглядит тёмным

Для изображений, приведённых на рисунке 96, $W(v, L) \approx 16$ раз. Снижение средней мощности (частоты спайков) передаваемых в мозг сигналов от глаза, при свёртке сетчаточного изображения с весовыми функциями РП, является важным эффектом обеспечения экономии организмом энергии. Возможно другим, не менее важным эффектом, является выделение при свёртке контуров исходного изображения для

распознавания зрительных образов, как это видно на рисунке 95 в. Кроме того, контурный характер свёрнутого изображения совместно с движением глаз обеспечивает топографичность проекций сигнальных изображений в зрительной системе в процессе обучения зрению.

5.12. Цветовая константность

Поскольку мы принимаем гипотезу о деконволюции (восстановлении) свёрнутого изображения сетчатки глаза в некоторой зоне мозга и наличии здесь устойчивого цветового шаблона, то мы можем предполагать возможность расслоения восстановленного изображения на цветовые компоненты. Эти компоненты могут быть субдискретизованы интерполяцией, как показано на рисунке 93 в. Кроме этого, при наличии цветовых компонент возникает возможность управления ими путём синхронного изменения величины их сигналов. Простые сигналы управления можно получить с периферии сетчатки глаза. Их простота обусловлена тем, что периферия сетчатки освещена равномерно, благодаря особенности оптической схемы глаза, открытой А. Л. Ярбусом [46]. Сигналы всего трёх колбочек различных видов L , M , и S , расположенных на периферии сетчатки, могут дать три сигнала об уровне адаптации этих колбочек к средней освещенности сетчатки глаза. На рисунке 97 приведены характеристики реакции зрительной системы на свет при дневном освещении, полученные по данным [56], в логарифмическом масштабе, как по яркостной

горизонтальной оси, так и вертикальной оси реакции. В таком представлении хорошо видны большие линейные участки как динамических (при различных уровнях адаптации), так и статической (уровня адаптации) характеристик при их монотонном возрастании.

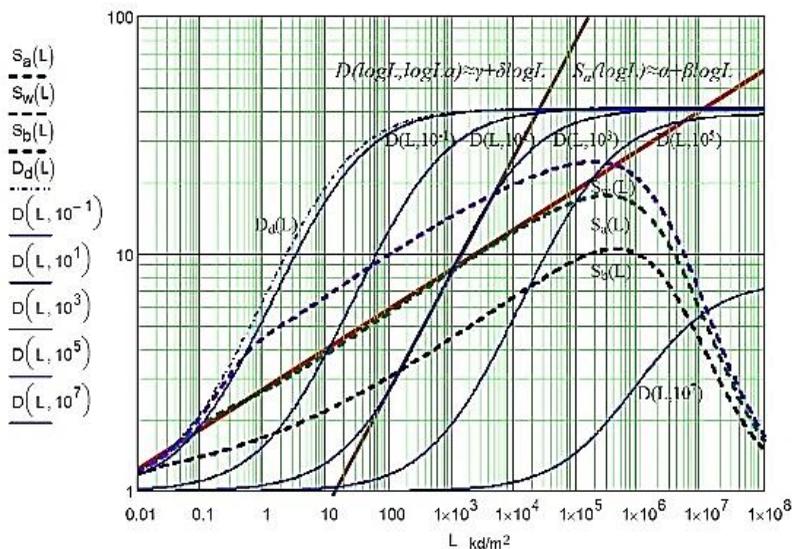


Рисунок 97 – Характеристики реакции зрительной системы на свет, при дневном (колбочковом) зрении: $D(L, 10^{-1}) - D(L, 10^7)$ – динамические, при различных уровнях адаптации, в том числе – $D_d(L)$ – при темновой адаптации, $S_a(L)$ – статическая, $S_w(L)$ и $S_b(L)$ – уровней белого и чёрного, соответственно

Заметим, что ранее мы не упоминали о логарифмическом характере биологических реакций, протекающих в соответствии с законами Вебера–Фехнера и Стивенса, поскольку это не влияет на принцип моделирования описываемых процессов.

Способ управления цветовыми компонентами поясним на простом примере. Если статический сигнал периферийной L -колбочки превышает остальные, то все сигналы красной компоненты следует пропорционально уменьшить, уравнивая их со средними сигналами остальных компонент. Такое управление приведёт к тому, что изменение спектра освещения всего того, что попадает в поле зрения, не будет ощущаться человеком. Дневное освещение и освещение лампами накаливания не будет вносить

сильного изменения в ощущение цветности объектов наблюдения. Будет соблюдаться цветовая константность. В технике подобная операция называется автоматическим балансом белого. Усреднение возможно и суммированием значений элементов цветовых компонент, но такая операция сложнее по сравнению с оптическим усреднением. Оптическое усреднение приводит к тому, что значение граничных элементов сетчаточного изображения становится одинаковым, и для деконволюции сигнального изображения нет необходимости в передаче каждого граничного элемента, что дополнительно снижает среднюю мощность передаваемых сигналов изображения.

После деконволюции изображения в коре мозга и обеспечения нормального восприятия окружающей среды в полноте тональных и цветовых раскрасок актуальным становится вопрос о разрешающей способности зрительной системы и общем числе её элементов разрешения.

5.13. Разрешающая способность зрительной системы

Для фовеальной области существенным показателем качества является минимальный угол Δ разрешения зрительной системы. Он зависит от ряда факторов, в частности: качества фокусировки оптического изображения на сетчатке, размеров колбочек и рецептивных полей ганглиозных нейронов сетчатки, дрейфа глаз, формы колбочек. Причём угловой размер колбочек не является определяющим. Угловое разрешение Δ нормального глаза принято в среднем считать равным одной угловой минуте. При этом остроту зрения можно оценивать величиной: $V = 1'/\text{MAR}$, где: MAR (*Minimal Angle of Resolution*) – минимальный угол разрешения испытуемого при наблюдении опто типов (испытательных изображений), выраженный в угловых минутах [57]. То есть нормальная острота зрения равна единице. Заметим, что угловой размер колбочек в фовеальной области сетчатки оценивается [6] величиной 0,5 угловой минуты. Однако у некоторых людей острота зрения может быть равной даже четырём, что можно объяснить лишь учитывая динамическую природу зрения и неравномерную чувствительность колбочки в её поперечном сечении. При движении (дрейфе и треморе) глазного яблока каждая колбочка сканирует небольшой участок оптического изображения, причём эффективная

апертура колбочки может быть меньше её полного диаметра, чему, по-видимому, соответствует коническая форма светочувствительной части колбочки. Действительно, ближе к оси конуса путь пробега света наибольший, как и, по-видимому, сигнал колбочки. При сканировании меньшей апертурой чёткость сигнального изображения возрастает, а сканирование – движение глаз – необходимо также для противодействия их быстрой световой адаптации, приводящей в случае неподвижного оптического изображения на сетчатке к исчезновению сигнального изображения и появлению на одну–две секунды «пустого поля» [8].

5.14. Число элементов разрешения глаза

С информационной точки зрения важным является вопрос об оценке числа элементов разрешения зрительной системы. По точкам пересечения линий равной остроты зрения (см. рисунок 87) с горизонталью, проходящей через центр поля зрения, построим методом линейной интерполяции графики остроты зрения рисунок 98 *a* (слева от центра $V_l(\varphi)$ – красная линия), справа от центра $V_r(\varphi)$ – синяя линия). Аналогично, на рисунке 98 *б* показаны графики остроты зрения для точек пересечения с вертикалью вниз от центра $V_d(\theta)$ (красный) и вверх от центра $V_i(\theta)$ (синий). Найдем максимальные взвешенные по остроте зрения углы поля зрения: влево – Ψ_{lM} , вправо – Ψ_{rM} , вниз – Θ_{dM} , и вверх – Θ_{iM} в соответствии с выражениями:

$$\Psi_{lM} = \int_0^{\varphi_{lM}} V_l(\varphi) d\varphi, \quad (89)$$

$$\Psi_{rM} = \int_0^{\varphi_{rM}} V_r(\varphi) d\varphi, \quad (90)$$

$$\Theta_{dM} = \int_0^{\theta_{dM}} V_d(\theta) d\theta. \quad (91)$$

$$\Theta_{iM} = \int_0^{\theta_{iM}} V_i(\theta) d\theta, \quad (92)$$

где φ_{lM} , φ_{rM} , θ_{dM} , θ_{iM} – максимальные углы поля зрения влево, вправо, вниз, вверх, соответственно.

Численный расчёт даёт следующие значения углов в градусах:

$$\Psi_{lM} = 9,1^\circ; \quad \Psi_{rM} = 7,7^\circ; \quad \Theta_{dM} = 6,0^\circ; \quad \Theta_{iM} = 5,6^\circ. \quad (93)$$

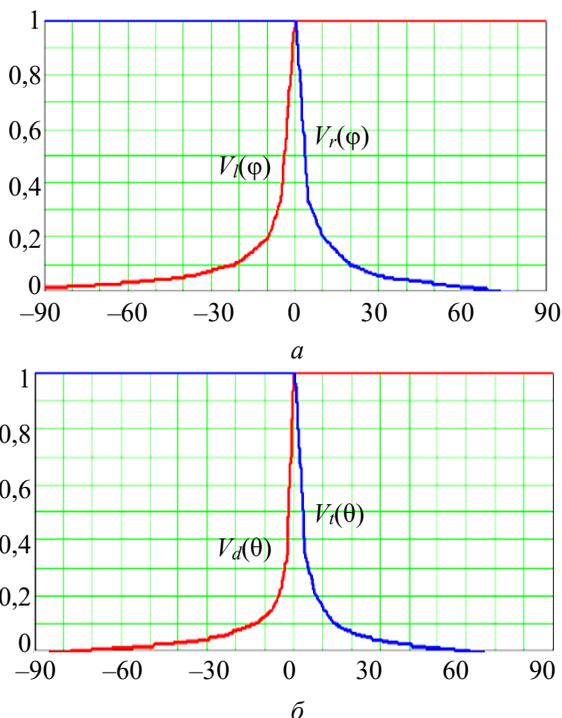


Рисунок 98 – График остроты зрения: по горизонтали – *a*; по вертикали – *b*

Поле зрения глаза, обозначенное на рисунке 87 штрих-пунктирной линией приблизительно имеет форму эллипса. Будем считать, что $a = \frac{\Psi_{lM} + \Psi_{rM}}{2\Delta}$ и $b = \frac{\Theta_{dM} + \Theta_{tM}}{2\Delta}$ приблизительно являются большой и малой полуосями этого эллипса. Выражая значение углового элемента разрешения глаза Δ в градусах, найдем площадь этого эллипса, которая является оценкой числа элементов разрешения глаза $N_o = \pi ab$. При нормальном зрении $\Delta = 1' = (1/60)^\circ$, а число элементов разрешения глаза $N_o = 550000$. Отметим, что это число в два раза меньше числа волокон зрительного нерва.

5.15. *Отображение: «глаз – зрительная кора»*

При отображении поля зрения (сетчатки глаза) на первичную зрительную кору все различные по величине элементы разрешения поля зрения отображаются в одинаковые по размеру элементы коры. На рисунке 99 *a* показана левая половина зрительной коры мозга,

состоящая из эллиптической по форме зоны первичной зрительной коры ($V1$), а также, окружающие её зоны вторичной ($V2$) и третичной ($V3$) коры. На зрительную кору наложена сетка полярной системы координат, соответствующая сетке полярной системы координат зрительного поля правого (контрлатерального) глаза (рисунок 99 б). Видно, что при отображении координатная сетка претерпевает сильные искажения. На рисунке 99 в приведен фрагмент дискретной модели первичной зрительной коры ($V1$), где обозначены оси косоугольной системы координат i и j . На рисунке 99 г приведен фрагмент дискретной модели половины поля зрения глаза в приведённых к коре элементах разрешения.

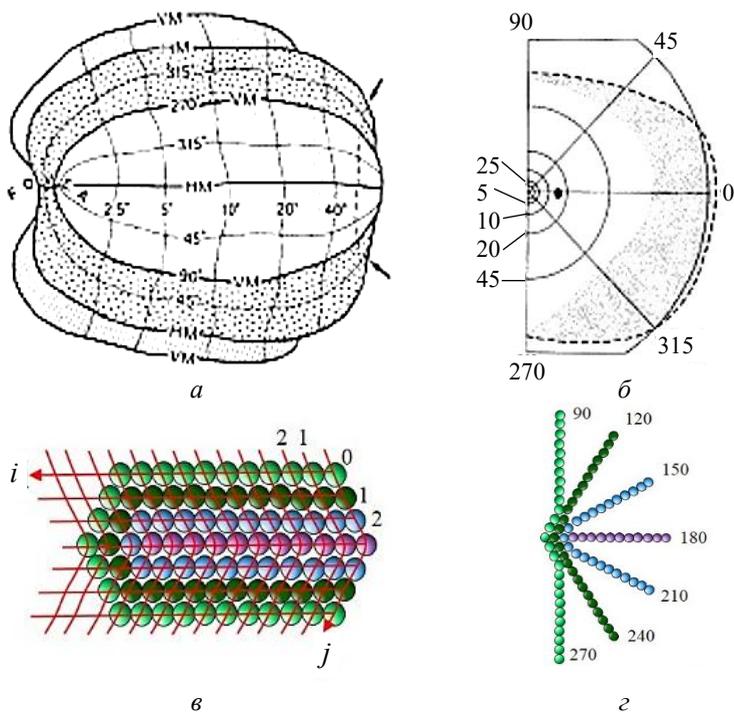


Рисунок 99 – Зрительная кора ($V1$, $V2$, $V3$) – а; половина поля зрения глаза – б; фрагмент дискретной модели первичной зрительной коры ($V1$) – в; фрагмент дискретной модели половины поля зрения глаза в приведённых к коре элементах разрешения – г

Очевидно, что переход от рисунка 99 г к рисунку 99 в можно произвести сворачиванием веера элементов, сводя его лучи в плотную упаковку.

6. Математическая модель улитки внутреннего уха

Поскольку разработка объёмного телевидения предполагает решение не только задач, связанных с объёмным изображением, но и задач, связанных с объёмным звуком, инженерам следует обратить особое внимание на получение их хорошего представления – математических моделей, описывающих работу слуховой системы человека. Важной частью слуховой системы является улитка внутреннего уха. Её модель предлагается читателю для ознакомления. Заметим, что этот раздел книги является, пожалуй, наиболее сложным для понимания в математическом отношении, но труд, затраченный на понимание, стоит того, поскольку речь идёт о нашем с Вами слухе.

Улитка внутреннего уха является составной частью слуховой системы человека. В ней осуществляется преобразование звукового сигнала в сигнал возбуждения волокон слухового нерва. Механизм этого преобразования в целом не имеет сколько-нибудь близких аналогов в технике, созданной человеком, что осложняет понимание его работы. Можно разбить процесс преобразования на отдельные фазы, каждая из которых имела бы знакомые нам аналоги. Такое разбиение представляет собой моделирование работы улитки внутреннего уха. Модель должна содержать не только качественное, но и количественное описание моделируемого объекта, что позволило бы как планировать эксперименты по дальнейшему его изучению, так и оценивать их результаты, совершенствуя саму модель.

Моделированию работы улитки посвящено множество работ теоретического и экспериментального характера, например, [58]–[62]. К сожалению, теоретические работы либо недостаточно прозрачны в отношении их физического смысла, либо не доведены до простых формул, позволяющих проводить расчёт модели с использованием современных универсальных систем компьютерной математики (пакетов *MathCAD*, *Matlab* и др.). По этим причинам работа с такими моделями затруднена.

Предлагаемая модель в удобном виде описывает изменение скорости распространения и затухания звуковой волны в прямом (вестибулярном) канале улитки, а также движение основной (базиллярной) мембраны. В заключительной части раздела высказывается предположение о трансверсальном, с точки зрения теории обработки сигналов, характере фильтрации (обнаружения) звуковых сигналов слуховой системой человека.

6.1. Улитка

Улитка представляет собой свёрнутую в два с половиной витка спирали костяную трубку, срез которой показан на рисунке 100. Трубка разделена внутри двумя перегородками (основной и рейснеровой мембранами) на три канала, по которым распространяется звуковая волна. В промежуточном канале

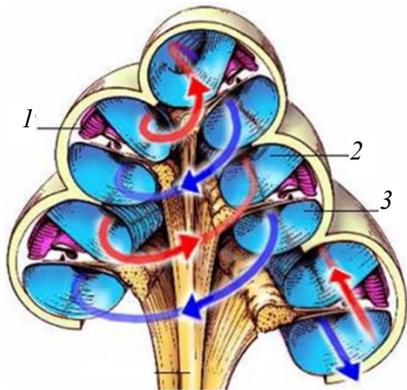


Рисунок 100 – Срез улитки внутреннего уха: промежуточный канал с органом Корти и покровной мембраной – 1; прямой канал – 2; обратный канал – 3; стрелками показаны прямой и обратный путь звуковой волны

находится орган Корти, прикреплённый к основной мембране и касающийся ещё одной – покровной мембраны (лучше сказать – мантии), закреплённой только с одного, ближнего к оси улитки, края. Движение этих мембран под действием звука преобразуется в слуховые сигналы нервных волокон с помощью волосковых клеток органа Корти.

Звуковой сигнал через систему косточек среднего уха поступает на гибкую мембрану овального окна прямого канала, изображенного на развёртке улитки, рисунок 101.

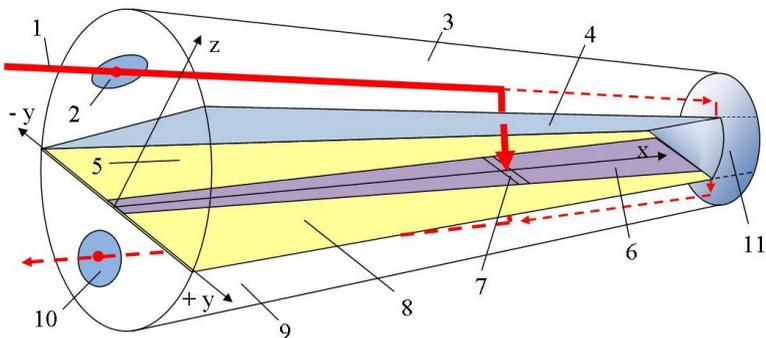


Рисунок 101 – Развёртка улитки (без органа Корти и покровной мембраны): 1 – звуковой сигнал; 2 – овальное окно; 3 – прямой канал; 4 – рейснерова мембрана; 5 – промежуточный канал; 6 – основная мембрана; 7 – одна из струн; 8 – костная оправа; 9 – обратный канал; 10 – круглое окно; 11 – геликотрема

Прямой канал заканчивается в вершине улитки отверстием (геликотремой), соединённым с началом обратного канала, ведущего к мембране круглого окна, выходящего в среднее ухо. Оба канала заполнены жидкостью – перилимфой, а находящийся между ними изолированный промежуточный канал заполнен жидкостью другого состава – эндолимфой. Основная мембрана армирована множеством жестких поперечных коллагеновых волокон – струн, натянутых на костную оправу поперек направления каналов. Число струн достигает 24 тысяч. Они настроены на различные звуковые частоты, высокие до 20 кГц у входа (выхода) прямого (обратного) канала и низкие до 20 Гц у геликотремы.

Звуковая волна определённой частоты, поступившая в прямой канал, доходит до зоны основной мембраны, резонирующей на эту частоту, и поглощается здесь органом Корти. Волны частотой меньше 20 Гц проходят через геликотрему в обратный канал, заставляя колебаться мембрану круглого окна. Волны, не до конца поглощенные органом Корти, также доходят до круглого окна обратного канала через основную мембрану или геликотрему.

Тонкая и податливая мембрана Рейснера легко передаёт звук любой частоты из основного канала в промежуточный, и поэтому может не учитываться при гидромеханических расчётах.

Ширина основной мембраны изменяется, увеличиваясь по направлению к геликотреме. На рисунке 102 изображен небольшой по длине (величиной dx) участок мембраны (условной струны).

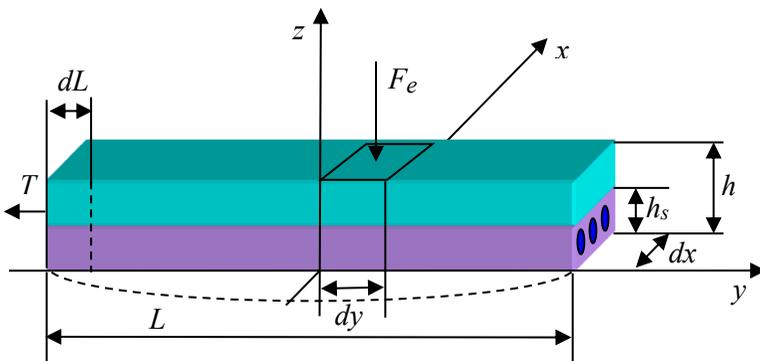


Рисунок 102 – Элемент основной мембраны – условная струна

Толщина h и длина L участка являются функциями его удалённости от начала узкого конца мембраны. Упругая часть участка основной мембраны толщиной h_s , состоящая из коллагеновых волокон и соединительной ткани, изображена на рисунке нижним слоем, а верхний слой соответствует нагрузке, которую создает прикрепленный к мембране орган Корти. Под воздействием переменного давления основная мембрана и её участок изгибаются, как показано на рисунке дугообразной пунктирной линией в одну, а затем в другую сторону.

6.2. Струна

Будем представлять основную мембрану улитки как совокупность натянутых струн подобно ряду струн фортепьяно или арфы. В отличие от музыкальных инструментов струны здесь соединены общей перепонкой, которая весьма слабо связывает их между собой, но добавляет сопротивления колебаниям со стороны жидкостей промежуточного и обратного каналов улитки, в которых струны колеблются. Струны имеют дополнительную нагрузку в виде органа Корти, что имеет свой аналог в музыкальных инструментах, где стальные струны обматываются медной проволокой для понижения частоты их звучания.

Одна из таких струн (рисунок 102) может содержать несколько коллагеновых волокон, которые вместе натянуты со средней силой T , приложенной к сечению струны $h_s \times dx$ (где dx – ширина струны; h_s – толщина упругого слоя струны, зависящая от x). Эту силу можно определить выражением:

$$T = \varepsilon E h_s dx \quad (94)$$

где E – модуль Юнга, коэффициент, характеризующий упругость материала струны, $\varepsilon = dL/L$ – относительное удлинение струны под действием силы T .

Рассмотрим теперь силы, действующие на элемент струны. На поверхность элемента действует сила:

$$F_e = P \sin(\Omega t) dx dy \quad (95)$$

где P – амплитуда давления звуковой волны; $\Omega = 2\pi f$ – круговая частота звука; f – частота звука; t – время; dx , dy – ширина и длина элемента струны.

Сила инерции элемента, согласно второму закону Ньютона, равна:

$$I_e = \rho h dx dy \frac{\partial^2 z}{\partial t^2}, \quad (96)$$

где $\rho h dx dy$ – масса элемента; $(\partial^2 z / \partial t^2)$ – ускорение элемента; ρ – средняя плотность элемента струны (двухслойного); h – толщина струны.

При движении элемент испытывает силу сопротивления жидкости окружающей среды, направленную противоположно его скорости:

$$R_e = \alpha \frac{\partial z}{\partial t} dx dy, \quad (97)$$

где α – коэффициент вязкости среды; $(\partial z / \partial t)$ – скорость элемента.

Сила упругости струны, действующая на элемент и стремящаяся вернуть его в положение равновесия, зависит от искривления струны в месте расположения элемента струны и в линейном приближении равна:

$$U_e = T \frac{\partial^2 z}{\partial y^2} dy = \varepsilon E h_s \frac{\partial^2 z}{\partial y^2} dx dy. \quad (98)$$

Баланс всех сил, действующих на элемент струны (уравнение его движения) определяется соотношением:

$$R_e + I_e - U_e = F_e \quad (99)$$

Подставляя в (99) выражения для составляющих сил, получим дифференциальное уравнение вынужденных колебаний струны в вязкой среде [63]:

$$\alpha \frac{\partial z}{\partial t} + \rho h \frac{\partial^2 z}{\partial t^2} - \varepsilon E h_s \frac{\partial^2 z}{\partial y^2} = P \sin(\Omega t). \quad (100)$$

Вводя обозначения:

$$\beta = \frac{\alpha}{\rho h}; \quad \eta^2 = \frac{\varepsilon E h_s}{\rho h}; \quad \chi = \frac{P}{\rho h},$$

запишем его в виде:

$$\beta \frac{\partial z}{\partial t} + \frac{\partial^2 z}{\partial t^2} - \eta^2 \frac{\partial^2 z}{\partial y^2} = \chi \sin(\Omega t). \quad (101)$$

Найдем решение этого уравнения $z(y, t)$ для установившихся стационарных колебаний, при которых вид решения не зависит от начальных условий (формы струны при $t = 0$). Этот случай соответствует периодическому звуковому сигналу и является весьма

характерным, хотя следует учитывать, что, например, при анализе речи сигнал можно считать стационарным в течение всего 10 миллисекунд. Граничные условия для концов струны определим равенством: $z(y, t) = 0$, при $y = \pm L/2$, что соответствует их жёсткому закреплению.

Кроме периодичности (во времени) звукового сигнала сделаем предположение о периодичности колебаний струны по пространственной координате y , считая струну бесконечной по длине. Такой чисто математический приём позволяет найти решение в классе периодических функций, а затем ограничиться реальным интервалом колебания (длиной) струны. Периодическую функцию можно в общем случае представить рядом Фурье. Учитывая приведённое выше граничное условие и при учёте движения струны под воздействием внешней силы, получим следующее представление вида искомого решения для коэффициентов разложения Фурье $c_n(t) = a_n \sin(\Omega t - \varphi_n)$:

$$z(y, t) = \sum_{n=1}^{\infty} c_n(t) \cos \frac{n\pi y}{L} = \sum_{n=1}^{\infty} \sin(\Omega t - \varphi_n) \cos \frac{n\pi y}{L}, \quad (102)$$

Амплитуду a_n и нулевую фазу φ_n каждой пространственной гармоники с номером n необходимо определить в процессе решения уравнения (101). Нулевая фаза колебаний гармоники определяется при сравнении с фазой звуковых колебаний в районе струны.

По всей длине реальной (ограниченной) струны на все её элементы действует в любой момент времени одинаковое давление звуковой волны, но для поддержания колебаний на условно бесконечной струне давление должно меняться вдоль неё на противофазное на каждом смежном по координате y участке длиной L . Последнее означает, что следует умножить коэффициент χ в уравнении (101) на функцию меандра $M(y)$ с периодом $2L$, изображённую на рисунке 103. Её можно разложить в ряд Фурье [64]:

$$M(y) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(1 - (-1)^n (-1)^{(n+3)/2})}{n} \cos \frac{n\pi y}{L} \quad (103)$$

Выражение (103) показывает, что чётные пространственные гармоники разложения равны нулю, а из нечётных гармоник каждая вторая берётся со знаком минус. На рисунке 103 приведены первая, третья и пятая гармоники разложения, причём третья гармоника инвертирована. Можно видеть, что для всех гармоник выполняются граничные условия, то есть их значения равны нулю при $y = \pm L/2$.

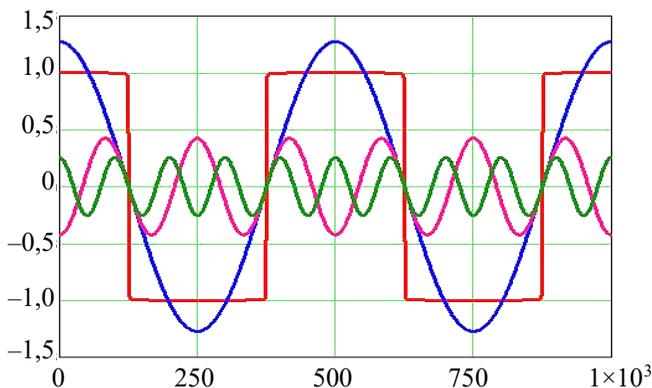


Рисунок 103 – Разложение функции меандра на гармонические составляющие

Отсутствие чётных пространственных гармоник в выражении для силы давления, вызывающей колебания струны, означает, что при случайном возникновении чётной гармоники в колебаниях струны она быстро исчезнет благодаря некомпенсируемым силам трения. Следует отметить, что имеются опытные данные, указывающие на наличие особых поперечных противофазных колебаний основной мембраны улитки, вызванные движением внешних волосковых клеток органа Корти. Это обстоятельство следует учитывать при задании длины струны L .

Запишем выражения для производных, входящих в (101), дифференцируя (102):

$$\frac{\partial z}{\partial t} = \sum_{n=1}^{\infty} \Omega a_n \cos(\Omega t - \varphi_n) \cos \frac{n\pi y}{L},$$

$$\frac{\partial^2 z}{\partial t^2} = \sum_{n=1}^{\infty} \Omega^2 a_n \sin(\Omega t - \varphi_n) \cos \frac{n\pi y}{L},$$

$$\frac{\partial^2 z}{\partial t^2} = \sum_{n=1}^{\infty} \frac{n^2 \pi^2}{L^2} a_n \sin(\Omega t - \varphi_n) \cos \frac{n\pi y}{L}.$$

Подставляя эти выражения в (101) с учётом (103), сокращая на величину $\cos(n\pi y/L)$ и обозначая $\omega = (\pi\eta)/L$, получим для каждого из n выражение:

$$\begin{aligned} & \beta \Omega a_n \cos(\Omega t - \varphi_n) + (n^2 \omega^2 - \Omega^2) a_n \sin(\Omega t - \varphi_n) = \\ & = \frac{2(1 - (-1)^n)(-1)^{(n+3)/2}}{n\pi} \chi \sin(\Omega t). \end{aligned} \quad (104)$$

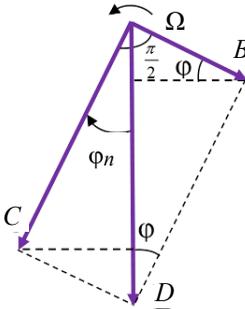


Рисунок 104 –
Соотношение векторов

Составляющие этого выражения можно представить векторами (рисунок 104), вращающимися против часовой стрелки с частотой Ω , с длинами:

$$B = \beta\Omega a_n, \quad C = (n^2\omega^2 - \Omega^2)a_n,$$

$$D = \frac{2(1 - (-1)^n)(-1)^{(n+3)/2}}{n\pi} \chi.$$

Очевидны следующие соотношения:

$$D_2 = B_2 + C_2, \quad \sin\varphi_n = B/D, \\ \cos\varphi_n = C/D, \quad \operatorname{tg}\varphi_n = B/C.$$

Из первого соотношения при подстановке значений входящих в него величин получим уравнение, решая которое найдем:

$$a_n = \frac{2(1 - (-1)^n)(-1)^{(n+3)/2} \chi}{n\pi \sqrt{(n^2\omega^2 - \Omega^2)^2 + \beta^2\Omega^2}}. \quad (105)$$

Из остальных соотношений с учётом значения a_n получим выражения для тригонометрических функций фазы φ_n :

$$\sin\varphi_n = \frac{\beta\Omega}{\sqrt{(n^2\omega^2 - \Omega^2)^2 + \beta^2\Omega^2}}. \quad (106)$$

$$\cos\varphi_n = \frac{n^2\omega^2 - \Omega^2}{\sqrt{(n^2\omega^2 - \Omega^2)^2 + \beta^2\Omega^2}}. \quad (107)$$

$$\operatorname{tg}\varphi_n = \frac{\beta\Omega}{n^2\omega^2 - \Omega^2}. \quad (108)$$

Учитывая, что:

$$z(y, t) = \sum_{n=1}^{\infty} a_n \sin(\Omega t - \varphi_n) \cos \frac{n\pi y}{L} = \\ = \sum_{n=1}^{\infty} a_n (\sin \Omega t \cos \varphi_n - \cos \Omega t \sin \varphi_n) \cos \frac{n\pi y}{L},$$

где $\sin(\Omega t - \varphi_n) = \sin \Omega t \cos \varphi_n - \cos \Omega t \sin \varphi_n$,

и подставляя в (102) значения величин из (105), (106), (107), получим решение уравнения (101):

$$z(y,t) = \sum_{n=1}^{\infty} \frac{2(1-(-1)^n)(-1)^{\frac{(n+3)}{2}} \chi \left[(n^2\omega^2 - \Omega^2) \sin \Omega t - \beta \Omega \cos \Omega t \right]}{n\pi \left[(n^2\omega^2 - \Omega^2) + \beta^2 \Omega^2 \right]} \cos \left(\frac{n\pi y}{L} \right), \quad (109)$$

в котором величины

$$\chi = P/\rho h, \quad \beta = \alpha/\rho h, \quad \omega = \pi\eta/L, \quad \eta^2 = \varepsilon E h_s / \rho h$$

являются в общем случае функциями координаты x и могут быть определены экспериментально.

6.3. Скорость звука

В разных средах звук распространяется с различными скоростями. В воздухе скорость звука 340 м/сек, в воде – 1400 м/сек, а в артерии человека скорость пульсовой волны составляет 10 м/сек.

Известно [65], что скорость звука в неограниченной среде определяется как:

$$c_s = 1/\sqrt{\rho_0 k_0},$$

где ρ – плотность среды, $k_0 = \Delta\rho_0/\rho_0\Delta p$ – сжимаемость среды, $\Delta\rho_0$ – изменение плотности среды при изменении звукового давления Δp .

Если звук распространяется в узкой трубке, то растяжимость её стенок под действием звукового давления эквивалентна увеличению сжимаемости среды, заполняющей трубку. Этим объясняется низкая скорость звука в артерии или резиновом огородном шланге. Растяжимость оценивается величиной $K = \Delta S/S\Delta p$, где S – площадь сечения трубки; ΔS – изменение площади сечения под действием звукового давления Δp .

Прямой канал улитки внутреннего уха (рисунок 100) расположен в кости черепа, которая обладает низкой растяжимостью, но стенки основного и, особенно, промежуточного канала, в который легко передаётся изменение звукового давления через рейснерову мембрану, выстланы более податливой клетчаткой, что способствует увеличению суммарной площади сечения каналов под действием давления звука. Однако наибольший вклад в увеличение площади сечения каналов вносит подвижность основной мембраны улитки. Представляя эту мембрану совокупностью струн, мы можем считать, что этот вклад обусловлен изгибом струны, соответствующим данному сечению.

Можно ввести эффективную сжимаемость жидкости каналов:

$$k = k_0 + K_w + K, \quad (110)$$

где K_w – растяжимость, обусловленная деформацией стенок каналов, K – растяжимость, обусловленная деформацией основной мембраны улитки. Учитывая два первых слагаемых выражения (110), получим начальную скорость звука в улитке у овального окна (без учёта действия основной мембраны):

$$c_0 = 1/\sqrt{\rho_0(k_0 + K_w)}. \quad (111)$$

Ранее отмечалось, что колебания струны содержат нечётные пространственные гармоники. При этом на длину струны L приходится n полувольт данной гармоники. Все, кроме одной полуволны, компенсируют изгиб друг друга. Нескомпенсированная полуволна вносит вклад в увеличение площади сечения звукопровода.

Если изменение давления $\Delta p = P$, то есть равно амплитуде звукового давления, то площадь, ограниченная одной полувольтной гармоникой с номером n в момент максимального звукового давления составит:

$$\Delta S_n = a_n \cos \varphi_n \int_{-L/2n}^{L/2n} \cos \frac{n\pi y}{L} dy = \frac{2La_n \cos \varphi_n}{n\pi}. \quad (112)$$

Используя это равенство, а также (105) и (107), получим выражение для растяжимости, обусловленной деформацией основной мембраны улитки:

$$K = \frac{1}{SP} \sum_{n=1,3}^{\infty} \Delta S_n = \sum_{n=1,3}^{\infty} \frac{8L(n^2\omega^2 - \Omega^2)}{n^2\pi^2 S\rho h((n^2\omega^2 - \Omega^2)^2 + \beta^2\Omega^2)}. \quad (113)$$

Скорость звука в основном (промежуточном) канале улитки:

$$c = \frac{1}{\sqrt{\rho_0 k}} = \frac{c_0}{\sqrt{1 + \rho_0 c_0^2 K}} = \frac{c_0}{\sqrt{1 + \rho_0 c_0^2 K U_-(K)}}. \quad (114)$$

В последнем выражении введена ступенчатая функция $U_-(K) = \begin{cases} 0, & K < 0 \\ 1, & K \geq 0 \end{cases}$, ограничивающая значение скорости звука

величиной c_0 для исключения слишком больших и даже мнимых значений скорости. Согласно выражению (114) скорость звука постепенно уменьшается от начальной c_0 до минимальной скорости, а затем растёт благодаря ускоряющему действию колеблющихся струн основной мембраны. Этот рост идёт до значений, которые могут значительно превышать c_0 . Реально скорость волны ограничена нелинейным характером ускоряющего действия, что не учитывается нашей линейной моделью. Отметим, что ускорение волны звука происходит в области улитки, где волна претерпела уже большое затухание и имеет пренебрежимо малую амплитуду.

6.4. Затухание звука

Распространяясь в какой либо среде, звук испытывает затухание, то есть амплитуда звуковой волны уменьшается. Затухание может быть обусловлено несколькими причинами: расхождением волны, рассеянием звука на препятствиях или его поглощением. В прямом канале улитки первые две причины незначительны, а поглощение происходит в результате перехода энергии звуковой волны в тепловую энергию при движении основной мембраны под действием силы звукового давления, из-за вязкости окружающей среды и трения органа Корти о покрывную мембрану.

При действии силы давления $F_e = P \sin(\Omega t) dx dy$ на элемент струны (рисунок 102) гармоническая составляющая колебаний струны с номером n , амплитудой a_n и фазой φ_n вносит свой вклад в смещение этого элемента за время dt . Этот вклад можно найти дифференцированием компоненты с номером n выражения (102) по времени:

$$dz_n(y, t) = a_n \Omega \cos(\Omega t - \varphi_n) \cos \frac{n\pi y}{L} dt. \quad (115)$$

Смещение элемента струны сопровождается возникновением силы трения. Часть силы давления на элемент F_e , пропорциональная множителю $\sin \varphi_n$ противодействует этой силе трения. Поскольку противодействующая сила отличается по своей фазе от фазы смещения элемента, её работа по смещению элемента уменьшается, причём также пропорционально множителю $\sin \varphi_n$.

Таким образом работа силы давления при смещении элемента струны на величину $dz_n(y, t)$ равна:

$$\begin{aligned} dA_n(y, t) &= F_e \sin^2(\varphi_n) dz_n(y, t) = \\ &= Pa_n \Omega \sin^2(\varphi_n) \cos(n\pi y/L) dx dy \sin(\Omega t - \varphi_n) \cos(\Omega t - \varphi_n) dt \end{aligned} \quad (116)$$

На длину струны L приходится нечётное число n полувольт гармоник. Чётное число разнополярных полувольт при интегрировании по длине струны даёт нулевое значение соответствующей части интеграла. Интегрируя выражение (116) по времени одного периода колебания струны $T_0 = 2\pi/\Omega$ и по длине одной полувольт гармоник (L/n) можно определить работу силы давления по преодолению сил трения при действии пространственной гармоник струны с номером n :

$$\begin{aligned} A_n &= Pa_n \Omega \sin^2(\varphi_n) dx \int_{-L/2n}^{L/2n} \cos \frac{n\pi y}{L} dy \int_0^{T_0} \sin(\Omega t) \cos(\Omega t - \varphi_n) dt = \\ &= \frac{2}{n} L Pa_n \sin^3(\varphi_n) dx. \end{aligned} \quad (117)$$

При действии всех гармоник работа против сил трения при вынужденных колебаниях струны в течении одного периода колебаний равна:

$$A = \sum_{n=1,3}^{\infty} A_n = 2PL dx \sum_{n=1,3}^{\infty} \frac{a_n}{n} \sin^3 \varphi_n. \quad (118)$$

Мощность потерь энергии звуковой волны на участке основной мембраны длиной dx и шириной L определяется выражением:

$$d\Phi = -PL\Omega \frac{dx}{\pi} \sum_{n=1,3}^{\infty} \frac{a_n}{n} \sin^3 \varphi_n. \quad (119)$$

Плотность потока мощности звуковой волны в жидкости составляет [68] $M = P^2/\rho_0 c$. Поток мощности в сечении основного и промежуточного каналов улитки

$$\Phi = P^2 S/\rho_0 c, \quad (120)$$

где S – площадь сечения.

Относительное уменьшение потока мощности на участке dx с учётом значений величин, входящих в (119) составит:

$$\frac{d\Phi}{\Phi} = -\frac{4\rho_0 c L \beta^3 \Omega^4 dx}{\pi^2 S \rho h} \sum_{n=1,3}^{\infty} \frac{(-1)^{(n+3)/2}}{n^2 [(n^2 \omega^2 - \Omega^2)^2 + \beta^2 \Omega^2]^2}. \quad (121)$$

Последнее выражение представляет собой дифференциальное уравнение с разделёнными переменными Φ и x . Интегрируя обе его части получим равенство:

$$\begin{aligned} \int \frac{d\Phi}{\Phi} &= \ln|\Phi| = \\ &= -\frac{4\rho_0 \beta^3 \Omega^4}{\pi^2 S \rho h} \int \sum_{n=1,3}^{\infty} \frac{(-1)^{(n+3)/2} c L dx}{n^2 [(n^2 \omega^2 - \Omega^2)^2 + \beta^2 \Omega^2]^2} + const. \end{aligned} \quad (122)$$

Определяя постоянную в правой части (122) из условия: $\Phi = \Phi_0 = \frac{P_0 S}{\rho_0 c_0}$, где P_0 – звуковое давление возле овального окна улитки при $x = 0$, получим:

$$\Phi = \Phi(x) = \Phi_0 \exp \left[-\frac{4\rho_0 \beta^3 \Omega^4}{\pi^2 S \rho h} \int_0^x \sum_{n=1,3}^{\infty} \frac{(-1)^{(n+3)/2} c L dx}{n^2 [(n^2 \omega^2 - \Omega^2)^2 + \beta^2 \Omega^2]^2} \right]. \quad (123)$$

Подставляя в (123) значения Φ и Φ_0 , получим зависимость амплитуды звукового давления в прямом (промежуточном) канале улитки как функцию расстояния от овального окна:

$$P = P(x) = P_0 \sqrt{\frac{c}{c_0}} \exp \left[-\frac{2\rho_0 \beta^3 \Omega^4}{\pi^2 S \rho h} \int_0^x \sum_{n=1,3}^{\infty} \frac{(-1)^{(n+3)/2} c L dx}{n^2 [(n^2 \omega^2 - \Omega^2)^2 + \beta^2 \Omega^2]^2} \right]. \quad (124)$$

Теперь мы имеем все необходимые зависимости для того, чтобы составить математическую модель улитки внутреннего уха.

6.5. Модель улитки внутреннего уха

На базе полученных в предыдущем разделе зависимостей можно составить математическую модель улитки внутреннего уха. Входные и выходные данные модели сведены в таблицу 2.

Таблица 2. Входные и выходные данные модели улитки внутреннего уха

№	Обозначение	Единица измерения	Наименование величины
1	α	кг/м ² сек	коэффициент вязкости
2	ρ_0	кг/м ³	средняя плотность мембраны*
3	ρ	кг/м ³	плотность звукопроводящей жидкости
4	c_0	м/сек	начальная скорость звука
5	f_0	сек ⁻¹	максимальная частота резонанса мембраны*
6	f_1	сек ⁻¹	минимальная частота резонанса мембраны*
7	n	–	номер пространственной гармоники
8	x_1	м	максимальное расстояние от начала мембраны *,**
9	x	м	расстояние от начала мембраны *,**
10	y	м	расстояние от середины струны мембраны *
11	z	м	смещение элемента мембраны *
12	L_0	м	минимальная ширина мембраны *
13	L_1	м	максимальная ширина мембраны *
14	P_0	Н/м ²	начальное звуковое давление
15	S	м ²	средняя площадь сечения звукопровода
16	F	сек ⁻¹	частота звука
17	$\beta(x)$	сек ⁻¹	коэффициент затухания
18	$\tau(x)$	сек	задержка распространения звуковой волны
19	$\omega(x)$	рад сек ⁻¹	круговая резонансная частота струны мембраны *
20	$c(x)$	м/сек	скорость звука
21	$h(x)$	м	толщина мембраны*
22	$L(x)$	м	ширина мембраны*
23	$P(x)$	Н/м ²	звуковое давление
24	$f(x)$	Н	резонансная частота струны мембраны*
25	$K(x)$	м ² /Н	растяжимость стенок звукопровода
26	$U(x)$	–	функция единичного скачка
27	$z(x, y, t)$	м	форма мембраны*

* – от основной (базиллярной) мембраны; ** – от овального окна

Некоторые входные данные являются функциями расстояния от начала основной мембраны улитки x . Другие взяты в своём среднем значении. В принципе, усредненные входные данные могут быть заменены на соответствующие зависимости от x или y , что лишь несколько увеличит компьютерное время расчёта параметров модели.

Составляющие модель зависимости могут использоваться как выходные данные модели, интересные для графического воспроизведения, так и в качестве входных данных, используемых для последующих этапов моделирования.

Можно использовать следующие выражения для величин, являющихся функциями расстояния от начала основной мембраны:

$$f = f(x) = f_0 \left(\frac{f_1}{f_0} \right)^{\frac{x}{x_1}}, \quad \omega = \omega(x) = 2\pi f(x),$$

$$h = h(x) = h_0 \left(\frac{h_1}{h_0} \right)^{\frac{x}{x_1}}, \quad \beta = \beta(x) = \alpha/\rho h(x),$$

$$L = L(x) = L_0 + (L_1 - L_0) \frac{x}{x_1}, \quad U_-[K(x)] = \begin{cases} 0, & K(x) < 0 \\ 1, & K(x) \geq 0 \end{cases}.$$

Тогда математическая модель улитки внутреннего уха описывается пятью соотношениями.

1. Изменение растяжимости стенок звукового канала, обусловленное деформацией основной мембраны улитки:

$$K(x) = \frac{8\rho_0 c_0^2}{\pi^2 S\rho} \sum_{n=1,3}^{\infty} \frac{L(x) [n^2 \omega^2(x) - \Omega^2]}{n^2 h(x) [(n^2 \omega^2(x) - \Omega^2)^2 + \beta^2(x) \Omega^2]}. \quad (125)$$

2. Изменение скорости распространения звуковой волны в жидкости над основной мембраной; скорость постепенно уменьшается от начальной c_0 до минимальной, а затем растёт, достигая значения c_0 :

$$c(x) = \frac{c_0}{\sqrt{1 + K(x)U_-[K(x)]}}. \quad (126)$$

3. Изменение амплитуды давления звуковой волны из-за потери её энергии при сопротивлении жидкости окружающей среды движению основной мембраны, а также из-за трения взаимодействия органа Корти и покровной мембраны улитки:

$$P(x) = P_0 \sqrt{\frac{c(x)}{c_0}} \exp \left[-\frac{2\rho_0 \beta^3 \Omega^4}{\pi^2 S \rho} \int_0^x \frac{c(x)L(x)}{h(x)} \sum_{n=1,3}^{\infty} \frac{(-1)^{(n+3)/2} dx}{n^2 [(n^2 \omega^2(x) - \Omega^2)^2 + \beta^2(x) \Omega^2]} \right]. \quad (127)$$

4. Изменение фазы колебаний основной мембраны с учётом задержки, которую испытывает звуковая волна при распространении в прямом (промежуточном) канале улитки:

$$\varphi(t, x) = \Omega \left(t - \int_0^x \frac{dx}{c(x)} \right), \quad (128)$$

5. Представление формы колеблющейся основной мембраны при распространении звуковой волны в прямом (промежуточном) канале улитки, соответствующей так называемой «бегущей волне Г. Бекеша», изображенной на рисунке 105 [58]:

$$z(t, x, y) = \frac{4P(x)}{\pi \rho h(x)} \times \sum_{n=1,3}^{\infty} \frac{(-1)^{\frac{(n+3)}{2}} \left[(n^2 \omega^2(x) - \Omega^2) \sin \varphi(t, x) - \beta(x) \Omega \cos \varphi(t, x) \right]}{n \left[(n^2 \omega^2(x) - \Omega^2)^2 + \beta^2(x) \Omega^2 \right]} \cos \left(\frac{n\pi y}{L(x)} \right). \quad (129)$$

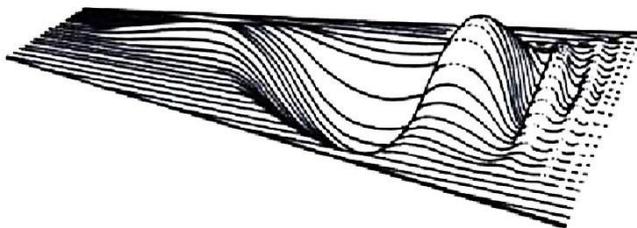


Рисунок 105 – Бегущая волна Г. Бекеша

В заключении данного раздела приведён пример расчёта предложенной модели волны Бекеша в системе *MathCAD*, где трёхмерный график волны был построен для звука частотой 1 кГц. На этом же графике построены две пространственные прямые, показывающие боковые границы основной мембраны. При использовании средств анимации (функции *FRAME MathCAD*), можно получить файл видеозаписи волны. Видеозапись можно наблюдать на сайте telenavt.ru в приложении к статье «Математическая модель улитки внутреннего уха».

6.6. Пути развитие модели улитки внутреннего уха

Предложенная линейная математическая модель улитки не затрагивает механизма работы органа Корти, сечение которого, соответствующее одной условной струне основной мембраны, показано на рисунке 106.

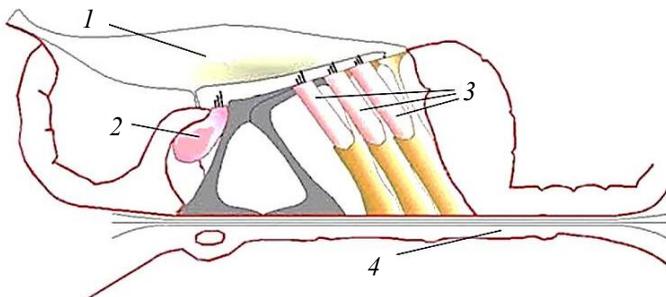


Рисунок 106 – Орган Корти: покровная мембрана (1); внутренняя волосковая клетка (2); внешние волосковые клетки (3); основная мембрана (4)

В предложенной модели действие органа Корти учитывается лишь выбором коэффициента $\beta(x) = \alpha/\rho h(x)$ путём произвольного задания величины α . Физиологические исследования показывают, что величина α задаётся центральной частью слуховой системы путём взаимодействия с внешними волосковыми клетками органа Корти, которые под воздействием эфферентных импульсов центральной нервной системы изменяют свою длину в такт с колебаниями основной мембраны. Сила механического действия этих клеток при слабом звуке компенсирует потери трения, возникающие при колебаниях основной и покровной мембран. В этом случае орган Корти работает как система с положительной обратной связью, причём прямую (афферентную) связь с центральной нервной системой обеспечивают как наружные, так и внутренние волосковые клетки. Если положительная обратная связь повышает коэффициент усиления слуховой системы, то отрицательная обратная связь при сильном звуковом сигнале может её снижать, для чего достаточно изменить фазу колебаний наружных волосковых клеток. Учёт работы механизма работы органа Корти позволит разработать приведённую в предыдущем разделе модель в части выбора коэффициента α для различных участков основной мембраны.

Ранее упоминалось о том, что улитка содержит около 24 тысяч коллагеновых упругих волокон, армирующих основную мембрану. Число внутренних волосковых клеток, расположенных в один ряд, достигает 3,5 тысяч, а число наружных, расположенных в три ряда, – 12 тысяч. Эти клетки равномерно распределены вдоль всей основной мембраны. Такое большое количество волосковых клеток позволяет предположить возможность существования у слуховой системы механизма трансверсальной фильтрации звуковых колебаний.

Трансверсальный [66] фильтр (см. рисунок 107) содержит линию задержки, состоящую из множества элементов τ_i с отводами и суммирующее устройство, осуществляющее так называемое взвешенное суммирование с заданными весами α_i .

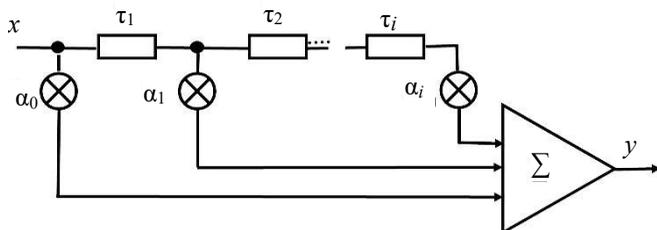


Рисунок 107 – Трансверсальный фильтр

Звуковая волна любой частоты воспринимаемого звукового диапазона, распространяясь в прямом (промежуточном) канале улитки, сначала плавно тормозится из-за действия колеблющейся основной мембраны, а затем резко ускоряется в районе резонанса. При торможении волна испытывает значительную задержку распространения, которая определяется по формуле

$$\tau_{1,2} = \int_{x_1}^{x_2} \frac{dx}{c(x)},$$

где x_1, x_2 – координаты участка задержки основной мембраны.

Аналогичная зависимость используется в правой части соотношения (128) нашей модели. Зная расстояние между соседними внутренними волосковыми клетками, мы можем определить элементарную задержку τ_i трансверсального фильтра улитки на данном участке основной мембраны. Суммирование сигналов различных волосковых клеток может осуществляться на

одном нейроне. Веса суммирования могут определяться запомненными в процессе обучения коэффициентами передачи синапсов дендритов этого нейрона.

Трансверсальные фильтры с конечной импульсной характеристикой (КИХ) позволяют легко осуществлять так называемую оптимальную фильтрацию входных сигналов. При этом с наибольшей эффективностью подавляются любые помехи за минимальное время приёма сигнала. В случае улитки внутреннего уха осуществляется хорошая развёртка примерно полутора периодов входного синусоидального сигнала (рисунок 105). За это время мозг может сделать заключение о наличии сигнала на входе слуховой системы. В отличие от резонансного фильтра, относящегося к классу фильтров с бесконечной импульсной характеристикой (БИХ), трансверсальный оптимальный фильтр согласуется с формой входного сигнала и не требует большого времени накопления одинаковых фрагментов входного сигнала. Например, при анализе речи время, в течение которого могут появляться такие фрагменты, оценивается величиной не более 10 миллисекунд. Таким образом, слуховая система, по-видимому, может использовать различные виды фильтрации: КИХ, БИХ, с независимой адаптивной регулируемой обратной связью, – что облегчает гибкое решение сложных задач распознавания слуховых образов.

Дополнение предложенной модели улитки возможными механизмами фильтрации позволит расширить сферу её применения и приблизиться к созданию математической модели слуховой системы человека. Создание такой модели может стимулировать изобретение и развитие новых технических средств воздействия на слуховой анализатор.

Наконец, приведём листинг программы Mathcad для построения волны Бекеша, анимацию которой можно видеть на сайте telenavt.ru, в приложении к статье «Математическая модель улитки внутреннего уха».

Листинг программы Mathcad для построения волны Бекеши

$x1:= 0$ $x2:= 30 \cdot 10^{-3}$ – начало и конец участка основной мембраны

$L1:= 0.06 \cdot 10^{-3}$ – ширина основной мембраны в начале,

$L2:= 0.2 \cdot 10^{-3}$ конце и

$L(x):= L1 + (L2 - L1) \cdot \frac{x}{x2}$ посреди участка (м)

$h1:= 0.04 \cdot 10^{-3}$ – толщина основной мембраны в начале,

$h2:= 0.04 \cdot 10^{-3}$ конце

$h(x):= h1 \cdot \left(\frac{h2}{h1} \right)^{\frac{x}{x2}}$ и посреди участка (м)

$f1:= 20 \cdot 10^3$ – резонансные частоты основной мембраны в начале,

$f2:= 20$ конце

$f(x):= f1 \cdot \left(\frac{f2}{f1} \right)^{\frac{x}{x2}}$ и посреди участка (m^{-1})

$\rho:= 2080$ – коэффициент вязкости среды, окружающей основную мембрану ($kg \cdot m^{-2} \cdot sec^{-1}$)

$\alpha:= 400$ – плотность основной мембраны ($kg \cdot m^{-2}$)

$\beta(x):= \frac{\alpha}{\rho \cdot h(x)}$ – коэффициент затухания основной мембраны (m^{-1})

$F:= 1 \cdot 10^3$ – частота звука (m^{-1})

$P0:= 1$ – начальная амплитуда давления звука ($N \cdot m^{-2}$)

$\Omega:= 2\pi \cdot F$, $\omega(x):= 2\pi \cdot f(x)$ – круговые частоты звука ($rad \cdot m^{-1}$)

$n := 1, 3 \dots 3$ – номер гармоники

$c_0 := 140$ – начальная скорость звука (м сек⁻¹)

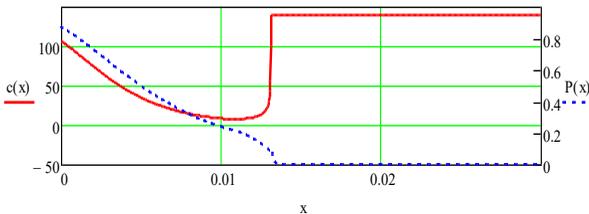
$\rho_0 := 1000$ – плотность среды распространения звука (кг м⁻²)

$S := 1 \cdot 10^{-6}$ – средняя площадь сечения звукопровода (м²)

Растяжимость стенок звукопровода из-за деформации основной мембраны (м² Н⁻¹)

$$K(x) := \frac{8c_0^2 \cdot L(x) \cdot \rho_0}{\pi^2 \cdot S \cdot \rho \cdot h(x)} \sum_n \frac{(n \cdot \omega(x))^2 - \Omega^2}{n^2 \left[\left[(n \cdot \omega(x))^2 - \Omega^2 \right]^2 + \beta(x)^2 \Omega^2 \right]}$$

$$c(x) := \frac{c_0}{\sqrt{1 + K(x) \cdot \Phi(K(x))}} \quad \text{– скорость звука (м сек}^{-1}\text{)}$$



$\Phi(x)$ – функция единичного скачка (Хевисайда)

Амплитуда звукового давления (Н м⁻²):

$$P(x) := P_0 \sqrt{\frac{c(x)}{c_0}} \exp \left[- \left(\frac{\rho_0 \beta(x)^3 \cdot \Omega^4}{\pi^2 S \cdot \rho \cdot h(x)} \right) \int_0^x \sum_n \frac{2(-1)^{\frac{n+3}{2}} c(x) \cdot L(x)}{n^2 \left[\left[(n \cdot \omega(x))^2 - \Omega^2 \right]^2 + \beta(x)^2 \Omega^2 \right]^2} dx \right]$$

$$\varphi(t, x) := \Omega \left(t - \int_0^x dx \frac{1}{c(x)} \right) \quad \text{– фаза колебаний основной мембраны (рад)}$$

$t := \frac{\text{FRAME}}{80 \cdot F}$ – функция задания времени при анимации (сек)

$t := \frac{0.3}{F}$ – функция задания времени без анимации (сек)

Отклонение основной мембраны от положения равновесия (мкм):

$$z(x,y) := \frac{4P(x)10^6}{\pi \cdot \rho \cdot h(x)} \sum_n \left[\frac{(-1)^{\frac{(n+3)}{2}}}{n} \cdot \frac{(n^2 \omega(x)^2 - \Omega^2) \sin(\varphi(t,x)) - \beta(x) \Omega \cos(\varphi(t,x))}{(n^2 \omega^2(x) - \Omega^2)^2 + \beta^2(x) \Omega^2} \right]$$

Вспомогательная функция ограничения на графике зоны, соответствующей основной мембране:

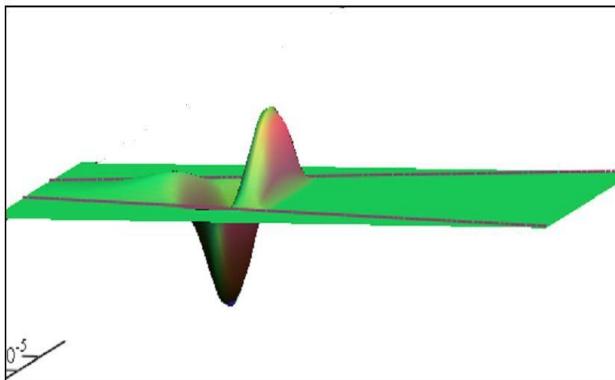
$$Z(x,y) := \text{if} \left(\left| y \right| < \frac{L(x)}{2}, z(x,y), 0 \right) \quad t0 := 0 \quad t1 := 3 \quad tg := 100$$

$$H(t) := \begin{pmatrix} 0.01 \cdot t \\ 0.000023t + 0.00003 \\ 0.001 \end{pmatrix}$$

$$G(t) := \begin{pmatrix} 0.01 \cdot t \\ -0.000023t - 0.00003 \\ 0.001 \end{pmatrix}$$

$$C := \text{CreateSpace}(H, t0, t1, tg)$$

$$D := \text{CreateSpace}(G, t0, t1, tg)$$



Z,D,C

ЗАКЛЮЧЕНИЕ

Телевидение (по-русски дальновидение) со временем приобретёт два новых свойства: «объёмность» и «интерактивность». Появление этих свойств в многообразных применениях телевидения улучшит их качество, подобно введению цвета или увеличению чёткости изображения. Новые свойства телевизионного изображения не должны ухудшать его качества и, по возможности, допускать использование имеющихся средств телевизионной связи. В частности, сетевая связь телевизионных систем, обладающая тенденцией к глобализации, реализуя принцип общедоступности информации, должна стать естественной средой для объёмного телевидения. Работая в области объёмного телевидения, я часто слышал вопросы: а зачем оно нужно – объёмное телевидение? Телевидение и так хорошее, даже слишком. Как-то профессор ЛЭТИ Р. Е. Быков, услышав про объёмное телевидение, рассмеялся: «Когда мы занимались цветным телевидением, нам часто твердили: "Зачем цвет, в чёрно-белом телевидении и так полно проблем"». Стоит сказать, что, во-первых, вопросы про объёмность были в основном со стороны простых телезрителей, но не специалистов в области прикладного телевидения, а, во-вторых, мой обычный ответ на такой вопрос со стороны молодого человека был: «моя задача дать возможность объёмности телевизионного изображения, а Вы придумаете, как её применить». Здесь, в конце книги, позволю себе включить фантазию и ответить так:

– "Объёмное телевидение даёт возможность:

- путешествовать в виртуальном времени, используя запись прошедших событий;
- превращаться в любых виртуальных существ, меняя свои аватары;
- трансформировать окружающую виртуальную среду, «материализовать» объекты и «трангрессировать» в выбранные точки мира телевизионного, подобно друзьям Гарри Поттера;
- ставить спектакли и снимать фильмы в виртуальных интерьерах;
- летать и прыгать как «Человек-паук» в виртуальном мире;
- использовать у себя дома различные тренажеры для овладения профессией летчика, моряка, шахтера, космонавта и даже хирурга;

- выполнять виртуальные тренировки и проводить виртуальные соревнования и игры в интерьерах знаменитых спортивных сооружений;
- спокойно смотреть 3D фильмы, снятые талантливыми кинематографистами, или встраиваться в сюжет фильма с разветвлённой системой продолжений в качестве зрителя или персонажа;
- создавать искусственные миры;
- становиться виртуальным туристом...”.

Но остановимся, поскольку фантазию, даже научную, в отличие от фантастики, следует держать на цепи – золотой цепи физического или серебряной цепи математического моделирования идеи. Каждая из перечисленных возможностей – целое направление со своими путями развития. Возьмем, например, виртуальный телетуризм. Представьте себе: сто тысяч аватаров телетуристов из разных стран на Красной площади, пришедших послушать бой курантов на Спасской башне. Если протокол виртуальной реальности запрещает аватарам пересекаться, то толкотня будет невообразимая. Технически можно снять запрет пересечений и превратить аватаров в призраки. Все ли пользователи это одобряют? Можно, при определённой плотности телетуристов сделать всех невидимыми и бесплотными оставив только физических туристов. Можно дать в руки телетуристам разнообразные фильтры – возможность проводить селекцию коллег по туризму по каким-либо признакам. Тут мы переходим к вопросам этики, эстетики и попадаем в сферу юриспруденции. В общем, жить в телевизионном мире не менее сложно и интересно, чем в физическом. Правда здесь мы можем при наличии нейро-интерфейса вершить всё по мановению мысли, но и остальные теленавты могут быть не хуже нас.

Глобальная система объёмного телевидения, при своем появлении, в очередной раз, но с бóльшей силой, покажет человеку, как мала и как разнообразна наша планета. Мы будем иметь в руках виртуальную копию нашего мира, с помощью которой каждый сможет переместиться в любое обжитое место этого мира и даже пообщаться с людьми в этом месте вне зависимости от их родного языка. Современного человека – потребителя, быстро адаптирующегося к любым технологическим новшествам, трудно удивить

чем-либо, но создатели этих новшеств пытаются удивить и при этом, с неизбежностью, ведут нас по пути познания себя. Мы – часть мира людей, и всё более совершенные системы связи между людьми постепенно рисуют в нашем сознании картину единства человечества. В этом единстве, как в зеркале, мы узнаем себя, таких умных и хороших, и таких несовершенных.

Познать себя – это кажется невозможным. Чем больше у нас знаний, тем больше эти знания меняют нас, и мы должны уже познавать себя бóльшего, чем прежде. Что движет нами в процессе познания? Ответ – интерес! Вот, например, интересно как мы видим и что такое видеть? Даже незрячий человек как-то «видит», благодаря наличию слуха, осязания и других чувств. Чувства помогают ему нарисовать в своем сознании «картину» окружающего мира и действовать в этом мире вполне адекватно используя свой интеллект. Не зная толком, как мы видим, мы уже взялись за создание *искусственного интеллекта*. Кажется, что смысл, который мы вкладываем в этот термин, несколько ограничен. Если бы попросить человека начала девятнадцатого века объяснить процесс видения, он вероятно для объяснения привлёк бы аналогию с процессом рисования картины. В конце XIX века появилась аналогия с процессом фотографии. В XX веке просвещённый интеллигент сравнит зрение с кинематографом и даже с телевидением. Все эти сравнения безусловно полезны, выявляя все новые свойства зрения, но и вредны, создавая у нас ложные представления о нём. Дело познания осложняется ещё и специализацией учёных, проявляющих интерес к зрению. Действительно, трудно представить себе биолога, хорошо владеющего техникой телевидения, а с другой стороны, что ожидать от «технаря», рассуждающего о физиологии зрения? А тут ещё проблема у той и другой стороны процесса познания со знанием специфической математики, необходимой для построения и чтения математических моделей, полноценно описывающих процессы зрения. Наконец, есть еще одна проблема, связанная с попыткой объяснить другому человеку то, что сам вроде бы понимаешь – проблема научения. Эта проблема возникает из-за того, что у каждого из нас своя картина мира, созданная при различиях жизненного опыта. Связать миры – сложная задача.

Развитие техники передачи изображений шло таким образом, что ни одно из её направлений не отражает в сколь-либо полном объёме передачу изображений зрительной системой человека [69]. Попробуем, тем не менее, описать в общих чертах её работу, по крайней мере, на её начальном (не интеллектуальном) участке, пользуясь, по возможности, технической терминологией. Изображение внешнего мира строится на сетчатке глаза с помощью перестраиваемой линзы – хрусталика. Сетчатка содержит три вида фоторецепторов: чувствительные (ночные) палочки и дневные колбочки трёх различных видов, чувствительные к красному, зелёному и синему областям спектра света. Фоторецепторы образуют первый слой сетчатки глаза. Второй слой сетчатки – биполярные клетки – соединены с фоторецепторами и либо прямо передают их сигнал к третьему слою ганглиозных клеток, либо передают его к этому слою инвертируя. Слой ганглиозных клеток сетчатки формирует своими аксонами (выходами) зрительный нерв. Важной особенностью сетчатки глаза является наличие в ней ещё двух дополнительных слоев клеток (нейронов): горизонтальных, между первым и вторым слоями, и амакриновых, между вторым и третьим слоями. Благодаря этим дополнительным слоям на каждый ганглиозный нейрон (волокно зрительного нерва) поступают сигналы сразу множества фоторецепторов. Эти фоторецепторы образуют рецептивное поле ганглиозного нейрона. Обычно это рецептивное поле, как бы состоящее из двух кругов (внешнего и внутреннего), имеет круговую коаксиальную форму, Эти круги находятся в оппонентных (противоположных) отношениях. Например, если в центральном малом круге сигналы фоторецепторов инвертированы, то в периферийном большом круге – не инвертированы. Такая организация рецептивного поля ганглиозного нейрона приводит к тому, что если нейрон «наблюдает» однородную по тону и цвету окраску объекта, то сигналы от обоих кругов равны, противоположны по знаку и компенсируют друг друга. То есть сигнал ганглиозного нейрона и его зрительного волокна равен нулю, и в мозг (зрительную кору) человека ничего не передаётся! Что это даёт? Это дает *экономия энергии, затрачиваемой на передачу сигнала* по длинному пути из глаза в зрительную кору, что биологически очень важно. Напомним, что у зрительного нерва около миллиона волокон. Что тогда передаётся

в мозг? Передаются сигналы от различных перепадов яркости объекта наблюдения в пределах одного рецептивного поля, когда сигналы от обоих его кругов не компенсируют друг друга. В итоге в первичную зрительную кору от глаза по зрительному нерву передается *контурное изображение* наблюдаемого объекта. На первый взгляд это кажется абсурдным и в книгах по физиологии зрения об этом умалчивается. Мы ведь видим никакое не контурное изображение, а сплошное чёрно-белое или цветное! В чём тут дело? Дело в том, что в мозге контурное изображение *восстанавливается* до полноценного, которое мы и воспринимаем. Можно ли провести такую операцию в принципе? Можно, для дискретного изображения, решая по определённому алгоритму так называемую задачу Дирихле. При этом восстановление полного (интегрального) изображения из контурного ведётся от границы изображения к центру быстро и непрерывно концентрическими расчётными волнами. Решение задачи Дирихле в данном случае имеет чёткую математическую основу, подробно описанную в настоящей книге. По сути, эта задача сводится к отысканию интеграла (интегрального изображения) определённой производной (контурного изображения). Ещё до интегрирования контурного изображения его можно использовать для *распознавания* заключённых в нём образов объектов по их контурной форме. Такова важная особенность зрения.

Известно много различных видов оппонентности рецептивных полей. Например, *красно – зелёная*, когда в центральном кругу преобладают инвертированные сигналы «красных» колбочек, а в большом периферийном – неинвертированные «зеленые» или *желто – синяя* с пространственным (коаксиальным) разделением, или из одинаковых кругов, наложенных друг на друга, или *чёрно – белая* и др. Цветовые оппонентности объясняют теорию четырёхцветового зрения Э. Геринга [72].

Восстановление интегрального изображения из контурного начинается с периферии сетчатки, которая, благодаря особой оптике глаза, открытой А. Л. Ярбусом [46], освещена равномерно усреднённым по спектру светом, попадающим в глаз. При таком восстановлении обеспечивается эффект *цветовой константности* зрения.

Ещё одной особенностью зрения является постоянное движение глаз в процессе зрения. Эта особенность представляется весьма загадочной многим физиологам зрения, в частности,

нобелевскому лауреату Д. Хьюбелу [73], изучавшему рецептивные поля зрительной коры [41]. По существу, в этом случае загадкой является механизм пространственной константности зрения, когда несмотря на движения глаз окружающий мир представляется неподвижным.

Движение глаз необходимо для исключения эффекта «пустого поля», описанного А. Л. Ярбусом [8], когда при исключении движения изображения по сетчатке глаза, уже через одну – две секунды видимое изображение исчезает, превращаясь в однородный серый фон. Такое превращение вызывается быстрой адаптацией фоторецепторов глаза к свету, а адаптация глазу необходима для перекрытия гигантского диапазона освещенностей окружающей среды в течение суток.

Пространственная константность зрения обусловлена многомодальностью и программным характером процесса зрения, когда мозг воспринимает окружающую среду посредством многих органов чувств, объединяя их показания и заранее планируя на какой угол повернуть глаза, рассматривая среду. Итог сводится к перемещению области ясного зрения по неподвижной «сетке» среды в процессе её рассматривания. По-видимому, восстановление интегрального изображения из контурного целесообразно производить на константном уровне зрительной системы.

Еще одной особенностью зрительной системы является отсутствие фоторецепторов (палочек) в зоне ясного зрения сетчатки, также в этой зоне отсутствуют «синие» колбочки. Тем не менее, мы не ощущаем ни серых пятен в этой области в сумерках, ни отсутствия в ней синего цвета днём. Здесь используется инерция накопления (суммирования) на «константном экране» ряда изображений с различными фазами движения глаз так, что если в одно мгновение в данном месте константного экрана отсутствует проекция синего, то в следующем мгновении из-за движения глаз она может появиться и сумма проекций окажется с «распространенным» синим. Аналогичный механизм может действовать для сумеречного (палочкового) зрения.

Отметим отсутствие – маскировку слепых пятен сетчаток глаз – мест, где сетчатку пронизывает пучок волокон зрительного нерва. Эта маскировка является следствием (особенностью) процесса интегрирования контурного изображения.

Перечислим отмеченные особенности работы зрительной системы:

- формирование контурного изображения (конволюция);
- восстановление интегрального изображения (деконволюция);
- «красно–зеленая» оппонентность зрения;
- «желто–синяя» оппонентность зрения;
- цветовая константность зрения;
- движение глаз, противостоящее адаптации;
- пространственная константность зрения;
- распространение синего на область ясного зрения;
- распространение чувствительности в сумерках на область ясного зрения;
- маскировка слепых пятен сетчаток глаз.

Большинство этих особенностей связано друг с другом, что создает известные сложности при изучении зрительной системы. Моделирование её работы, описанное в книге, позволяет понять назначение и возможный механизм реализации этих особенностей.

Зрительная система человека является дискретной системой, а как известно, дискретным системам формирования изображений присуща такая особенность, как алиасинг (биение пространственных частот) и появление муарных искажений. Наиболее ярко эти особенности проявляются при наблюдении особых тестовых изображения – колец Френеля. Поскольку вопросы метрологии часто играют ключевую роль при разработке новых телевизионных систем, в книгу введён раздел о кольцах Френеля.

На втором месте после зрения по уровню информационной насыщенности в современном мире стоит такое чувство человека, как слух. Создание эффективных систем объёмного звука для объёмного телевидения находится в прямой связи с пониманием работы соответствующего органа этого чувства. В конце книги рассмотрена работа приёмника звука – улитки внутреннего уха человека. Надеемся, что знания, полученные из книги, найдут практическое применение при создании систем объёмного телевидения.

Литература

1. Варгин, П. Объёмное телевидение. Позаботимся о здоровье глаз / П. Варгин // «625».1999. № 1.
2. Варгин, П. Объёмное телевидение – взгляд из России / П. Варгин // Broadcasting. 2007. № 6 (Ч. 1). С. 54–56. № 7 (Ч. 2). С. 34–39.
3. Варгин, П. Объёмное телевидение – символ XXI века / П. Варгин // Broadcasting. 2011. № 4 (Ч. 1). С. 36–39. № 5 (Ч. 2). С. 38–41. № 6 (Ч. 3). С. 36–38. № 7 (Ч. 4). С. 48–51. №8 (Ч. 5). С. 38–39. 2012. № 1/2 (Ч. 6). С. 35–37.
4. Объёмное телевидение. <https://ru.wikipedia.org/wiki/>.
5. Adelson E., Bergen J. The plenoptic function and the elements of early vision. M. Landy and J. A. Movshon, (eds) Computational Models of Visual Processing. 1991.
6. Хьюбел Д. Глаз, мозг, зрение; перевод с англ. О. В. Левашова, Г. А. Шараева; под ред. А. Л. Бызова. – Москва: – Мир, – 1990. – 240 с.
7. Раушенбах, Б. В. Пространственные построения в живописи: Очерк основных методов / Б. В. Раушенбах. – Москва: – Наука. – 1980. – 288 с.
8. Ярбус, А. Л. Роль движений глаз в процессе зрения / А. Л. Ярбус. – Москва: Наука, 1965. – 168 с.
9. Пацлаф, Р. Застывший взгляд. Физиологическое воздействие телевидения на развитие детей; пер. с нем. В. Бакусева / Райнер Пацлаф. – Москва: evidentis, 2003. –224 с.
10. Авторское свидетельство СССР. Способ определения пространственной формы объектов № 174185, заявка № 3015625 от 06.04.1981 г., опубликовано 27.05.2012 г. Бюл. № 15, SU 1 840 824 A1 / Варгин П. С.
11. Яне, Б. Цифровая обработка изображений / Б. Яне. – Москва: Техносфера, 2007. – 584 с.
12. Holliman N. S., Dodgson N. A., Gregg E., Favalora G. E. and Pockett L. Three-Dimensional Displays: A Review and Applications Analysis // IEEE Transactions on broadcasting 57(2), 2011. PP. 362–371.
13. Ketchpel, R. D. A Direct-View Three-Dimensional Display Tube // IEEE Transactions on Electronic Devices Vol 5. 1963. PP. 324–329.
14. Langhans K. FELIX 3D Display: An Interactive Tool for Volumetric Imaging, Proceedings of SPIE // The International Society for Optical Engineering, May 2002.
15. Леушина, Л. И. Зрительное пространственное восприятие / Л. И. Леушина. – Ленинград: Наука, 1978. – 176 с.

16. Сэломон, Д. Сжатие данных, изображений и звука / Д. Сэломон. – Москва: Техносфера, 2006. – 368 с.
17. Ричардсон, Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения: справочное издание / Я. Ричардсон; пер. с англ. В. В. Чепыжова. – Москва: Техносфера, 2005. – 366 с.
18. Poggio, G. F., Fischer, B. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey // *Journal of Neurophysiology*. 1977. 40(6). С. 1392–1405.
19. Arshavsky V. Y., Wensel T. G. Timing Is Everything: GTPase Regulation in Phototransduction // *Investigative Ophthalmology & Visual Science*. Vol.54, 2013.
20. van Hateren, J. H., Snippe, H. P. Simulating human cones from mid-mesopic up to high-photopic luminances // *J. Vision*, 7(4), 1, 1–11, 2007.
21. Тарасов, В. В. Инфракрасные системы «смотрящего» типа / В. В. Тарасов, Ю. Г. Якушенков. – Москва: – Логос, – 2004. – 444 с.
22. Michael A. Gerzon. Periphony: With-Height Sound Reproduction // *Journal of the Audio Engineering Society*, 1973, 21(1). PP. 2–10.
23. Варгин, П. С. Определение пространственной формы искусственных космических объектов методом телевизионной лазерной локации / П. С. Варгин, А. А. Умбиталиев, А. В. Черногубов // *Вопросы радиоэлектроники. Серия Техника телевидения*. – 2019. – Вып. 2. – С. 10–23.
24. Фотограмметрия: учебник / И. А. Краснощекова, О. Б. Нормандская, А. М. Кислова, В. В. Кислов. – Москва: Недра, 1978. – 471 с.
25. Лобанов, А. Н. Автоматизация фотограмметрических процессов / А. Н. Лобанов, И. Г. Журкин. – Москва: Недра, 1980. – 240 с.
26. Валюс, Н. А. Стереоскопия. Акад. Наук СССР. Комис. по науч. фотографии и кинематографии. – Москва: Изд-во Акад. наук СССР, 1962. – 379 с.
27. Грузевич, Ю. К. Оптико-электронные приборы ночного видения / Ю. К. Грузевич. – Москва: Физматлит, 2014. – 275 с.
28. Kawakita M., Kurita T., Hiroshi H. and Inoue S. HDTV Axi-vision Camera. // *Proc. IBC (International Broadcasting Convention) 2002, September, Amsterdam*. PP. 397–404. (2002).
29. Зубарев, Ю. Б. Видеоинформационные технологии объёмного и многоракурсного телевидения / Ю. Б. Зубарев, Ю. С. Сагдуллаев, Т. Ю. Сагдуллаев // *Вопросы радиоэлектроники. Серия Техника телевидения*. – 2011. – Вып. 1. – С. 32–44.

30. Аззам, Р. Эллипсометрия и поляризованный свет / Р. Аззам, Н. Башара; пер. с англ., под ред. А. В. Ржанова, К. К. Свиташева. – Москва: Мир, 1981. – 583 с.

31. Физика: Учебник / А. Т. Глазунов, О. Ф. Кабардин, А. Н. Малинин [и др.]; под ред. А. А. Пинского. – 4. изд. – Москва: Просвещение, 1999. – 431 с.

32. <https://vixion.jp/> Дата обращения 18.03.2025

33. Kuiper S., Hendriks B. H. W. Variable-focus liquid lens for miniature cameras // *Applied physics letters*, V 85, N 7. 2004.

34. Zhang W. J. B, Songjing Li. Progressive Multifocal Liquid Lenses Based on Asymmetric Freeform Surface Structure of Nonuniform Thickness Elastic Membranes with Different Constraints // *International Journal of Optics*, Issue 1 2032870. 2019.

35. Исаев, Ю. Н. Практика использования системы MathCad в расчетах электрических и магнитных цепей: учебное пособие / Исаев Ю. Н., Купцов А. М. – Москва: СОЛОН-Пресс, 2013. – 180 с.

36. Сайт *Tobii Technology, Inc* <https://www.tobii.com/>

37. Варгин, П. С. Испытательное изображение – кольца Френеля // *Вопросы радиоэлектроники. Серия Техника телевидения*. 2018. Вып. 3. С. 47–59.

38. Антипин М. В. Анализ и синтез телевизионной системы: Учебное пособие / М. В. Антипин; М-во высш. и сред. спец. образования РСФСР. Ленингр. электротехн. ин-т им. В. И. Ульянова (Ленина). – Ленинград: [б. и.], 1970. – 122 с.

39. Rafinazari M. Denoising and demosaicking of color images. Thesis for the Ph.D. degree. University of Ottawa, 2017.

40. Варгин, П. С. Методики измерения четкости и разрешающей способности цифровой телевизионной системы // *Вопросы радиоэлектроники. Серия Техника телевидения*. 2017. Вып. 4. С. 16–34.

41. Гонзалес, Р. Цифровая обработка изображений / Р. Гонзалес, Р. Вудс. – М.: Техносфера. 2006.

42. Марр, Д. Зрение: Информационный подход к изучению представления и обработке зрительных образов; пер. с англ. Н. Г. Гуревич; под ред. [и с предисл.] И. Б. Гуревича / Д. Марр. – Москва: Радио и связь, 1987. – 399.

43. Golay M. J. E. Hexagonal Pattern Transformation // *IEEE Trans. Computers*, C-18, 8, 733 – 740 (August 1969).

44. Middleton L., Sivaswamy J. Hexagonal Image Processing: A Practical Approach. Springer, 2005.

45. Schira M. M., Wade A. R., and Tyler C. W. Two-Dimensional Mapping of the Central and Parafoveal Visual Field to Human Visual Cortex // *J Neurophysiol* 97: 4284–4295, 2007.
46. Ярбус, А. Л. О работе зрительной системы человека. Адекватный зрительный стимул / А. Л. Ярбус // *Биофизика*. – 1975. – Т. 20. – № 5. – С. 916–920.
47. Шевелев, И. А. Волновые процессы в зрительной коре мозга / И. А. Шевелев // *Природа*. – 2001. № 12. – С. 1–13.
48. Баркалов, К. А. Образовательный комплекс «Параллельные численные методы»; методическое пособие. – Нижний Новгород: Изд-во Нижегородского госуниверситета им. Н. И. Лобачевского, 2011, 124 с.
49. Гиппенрейтер, Ю. Б. Движения человеческого глаза / Ю. Б. Гиппенрейтер. – Москва: Изд-во МГУ, 1978. – 256 с.
50. Соколов, Е. Н. Нейронные механизмы саккадических движений глаз. Институт психологии АН СССР // *Вопросы психологии*, – 1980, – 3, 1980. С. 64–77.
51. Юстова, Е. Н. Цветовые измерения (Колориметрия): научное издание; науч. ред. Ю. Н. Демков / Е. Н. Юстова. – Санкт-Петербург: Изд-во СПбГУ, 2000. – 398 с.
52. Williams D. R., MacLeod D, Hayhoe M. M. Foveal tritanopia. *Vis Res*. 1981; 21:1341–1356.
53. Steven M. LaValle. *Virtual reality*. Cambridge university press. 2020. 426 с.
54. Варгин П. С. Контурное кодирование изображений / П. С. Варгин // *Вопросы радиоэлектроники. Серия Техника телевидения*. – 2022. – Вып. 1. – С. 39–54.
55. Дейвис, Д. Онтогенез: от клетки до человека; пер. с англ. Т. Матешиной, Н. Ленцман / Джейми Дейвис. – Санкт-Петербург: Питер, 2017. – 349 с.
56. Hunt R. W. G. *The reproduction of color*. Sixth edition. John Wiley & Sons, Ltd, 2004. 702 с.
56. Ling-Qi Zhang, N. P. Cottaris, D H. Brainard. An Image Reconstruction Framework for Characterizing Early Vision. <https://doi.org/10.1101/2021.06.02.446829>.
57. Рожкова, Г. И. Зрение детей: проблемы оценки и функциональной коррекции / Г. И. Рожкова, С. Г. Матвеев. – Москва: Наука, 2007. – 314 с.
58. Georg Von Békésy. *Experiments in Hearing*. New York e.a., Mc.Grow-Hill Book Company, 1960, 745 p.
59. Бабич, В. М. О колебаниях базилярной мембраны во внутреннем ухе млекопитающих» / В. М. Бабич, С. М. Новоселова // В сб. *Записки научных*

семинаров ЛОМИ. Математические вопросы теории распространения волн. 1979. Т. 89. С. 54–62.

60. Allen J. Cochlear modeling. // IEEE ASSP Magazine. PP. 3–29, Jan. 1985.

61. Watts L. Cochlear Mechanics: Analysis and Analog VLSI. PhD thesis, California Institute of Technology. 1993.

62. Варин, В. П. Математическая модель слуховой улитки человека / Варин В. П., Петров А. Г. – Москва: Ин-т прикладной математики им. М. В. Келдыша РАН, 2009. – 26 с. (Препринт / Российская акад. наук, Ин-т прикладной математики им. М. В. Келдыша, № 96 за 2008 г.)

63. Крылов, А. Н. О некоторых дифференциальных уравнениях математической физики, имеющих приложение в технических вопросах; изд. 5-е. / А. Н. Крылов. – Москва-Ленинград: Гос. изд-во технико-теорет. лит-ры, 386 с.

64. Гоноровский, И. С. Радио-технические цепи и сигналы: учебник для вузов. – 4-е изд., перераб. и доп. / И. С. Гоноровский. – Москва: Радио и связь, 1986. – 512 с.

65. Исакович, М. А. Общая акустика / М. А. Исакович. – Москва: Наука, 1973. – 495 с

66. Рабинер, Л. Теория и применение цифровой обработки сигналов; перевод с англ. А.Л. Зайцева [и др.] ; под ред. Ю.Н. Александрова / Л. Рабинер, Б. Гоулд. – Москва: Мир, 1978. – 848 с.

67. Структурно-функциональная организация зрительной системы в обеспечении целенаправленной деятельности / Ю. Е. Шелепин, Е. Ю. Шелепин, В. М. Бондарко [и др.] // Успехи физиологических наук. – 2024. Т. 55, № 3. – С. № 3–21.

68. Твёрдотельные телекамеры: накопление качества информации / А. К. Цыцулин, Д. Ю. Адамов, А. А. Манцетов [и др.]. – Санкт-Петербург: СПбГЭТУ «ЛЭТИ», 2014. – 272с.

69. Kapustin V.V., Zahlebin A.S., Movchan A.K., Kuryachiy M.I., Krutikov M.V. Experimental assessment of the distance measurement accuracy using the active-pulse television measuring system and a digital terrain model // Computer Optics, Vol. 46(6), 2022.

70. Активно-импульсные телевизионные измерительные системы для обеспечения навигации транспортных средств в сложных метеоусловиях / В. В. Капустин, А. К. Мовчан, Е. В. Зайцева [и др.] // Транспортные системы и технологии. – 2018. Т. 4. № 1. – С. 68–83.

71. Hering E. Outlines of a Theory of the Light Sense. Cambridge, Mass: Harvard University Press. – 1964. – 317 с.

72. Martinez-Conde S., Macknik S. L. and Hubel D. H. THE role of fixational eye movements in visual perception. Neuroscience v 5, march, 2004.

Оглавление

Предисловие	3
От автора	6
Введение	8
1. Объёмное телевидение – символ XXI века?	12
1.1. Анализ термина «объёмное телевидение».....	13
1.2. Классификация систем объёмного телевидения.....	15
1.3. Схемы наблюдения.....	18
1.4. Когнитивная наука.....	21
1.5. Три мира.....	24
1.6. Захват сознания.....	29
1.7. Три части телевизионной системы.....	34
1.8. Зрительная система.....	50
1.9. Адаптация.....	57
1.10. Трансформации мира телевизионного.....	65
1.11. Объёмный звук.....	66
2. Телевизионный лазерный локатор	74
2.1. Способ телевизионной лазерной локации.....	75
2.2. Математическая модель ТВЛЛ.....	78
2.3. Однозональный вариант ТВЛЛ.....	79
2.4. Многозональный вариант ТВЛЛ.....	81
2.5. Импульсный лазер.....	82
2.6. Приём отражённого излучения лазера.....	83
2.7. Многоимпульсные варианты ТВЛЛ.....	85
2.8. Пространственное суммирование подсветки.....	87
3. Объёмный монитор открытого типа	89
3.1. Согласование конвергенции и аккомодации.....	90
3.2. Разрешение дисплея.....	96
3.3. Отслеживание глаз.....	97
3.4. Размытие вне фокуса.....	98
3.5. Динамический фовеальный рендеринг.....	99

4. Испытательное изображение – кольца Френеля.....	100
4.1. Кольца Френеля с синусоидальным профилем.....	103
4.2. Кольца Френеля с бинарным профилем.....	108
4.3. Сдвиги колец Френеля относительно растра изображения.....	111
5. Зрение.....	119
5.1. Контурное кодирование изображений.....	119
5.2. Выделение контура.....	120
5.3. Восстановление изображения.....	125
5.4. Гексагональная решетка.....	127
5.5. Восстановление исходного изображения.....	132
5.6. Сглаживание локальных дефектов изображения.....	136
5.7. Оппонентно-цветовое кодирование изображений.....	140
5.8. Дискретизация.....	148
5.9. Пространственная константность.....	150
5.10 Распространения синего цвета на область ясного зрения.....	151
5.11. Контурное изображение объектов наблюдения.....	153
5.12. Цветовая константность.....	166
5.13. Разрешающая способность зрительной системы.....	168
5.14. Число элементов разрешения глаза.....	169
5.15. Отображение: глаз – зрительная кора.....	170
6. Математическая модель улитки внутреннего уха...	172
6.1. Улитка.....	173
6.2. Струна.....	175
6.3. Скорость звука.....	180
6.4. Затухание звука.....	182
6.5. Модель улитки внутреннего уха.....	185
6.6. Пути развитие модели улитки внутреннего уха.....	188
Листинг программы Mathcad для построения волны Бекеша.....	191
Заключение.....	194
Литература.....	201

Научное издание

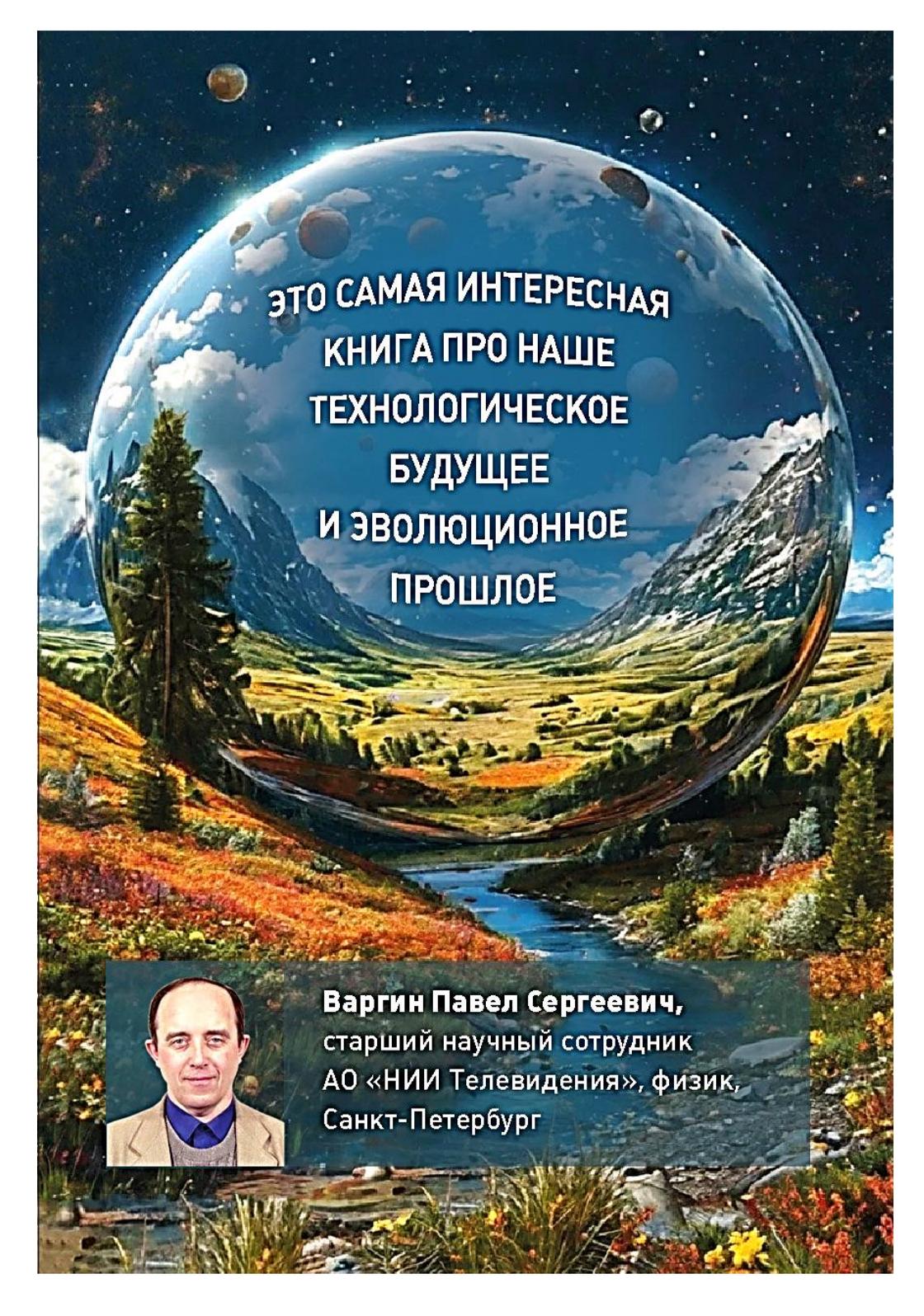
Павел Сергеевич Варгин

Объёмное телевидение и зрение

Редактор Л. П. Долгова

Подписано в печать 15.05.2025. Формат 60×84 1/16
Бумага офсетная. Печать цифровая. Печ. л. 15,0
Тираж 120 экз. Заказ 78.
Издательство АО «НИИ телевидения»
197121, Санкт-Петербург, Политехническая ул., 22

Отпечатано с готового оригинал-макета
Издательство СПбГЭТУ «ЛЭТИ»
197022, С.-Петербург, ул. Проф. Попова, 5Ф



**ЭТО САМАЯ ИНТЕРЕСНАЯ
КНИГА ПРО НАШЕ
ТЕХНОЛОГИЧЕСКОЕ
БУДУЩЕЕ
И ЭВОЛЮЦИОННОЕ
ПРОШЛОЕ**



Варгин Павел Сергеевич,
старший научный сотрудник
АО «НИИ Телевидения», физик,
Санкт-Петербург