**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ**

**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук

о.п. “Программная инженерия”

**РЕФЕРАТ**

**ТЕМА:** **«ЗРИТЕЛЬНЫЙ АНАЛИЗАТОР. ОБРАБОТКА ЗРИТЕЛЬНОГО СИГНАЛА.  РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ.»**

**Преподаватель**

Профессор департамента программной инженерии

факультета компьютерных наук

Харламов Александр Александрович

**Исполнитель**

Студент группы 205 Морозова А.В.

Москва 2020

**Оглавление**

[**1.** **Введение** 4](#_Toc59040576)

[**2.** **Зрительный анализатор** 6](#_Toc59040577)

[**Склера** 6](#_Toc59040578)

[**Роговица** 6](#_Toc59040579)

[**Хрусталик** 6](#_Toc59040580)

[**Радужная оболочка глаза** 7](#_Toc59040581)

[**Зрачок** 7](#_Toc59040582)

[**2.1** **Сетчатка** 7](#_Toc59040583)

[**Зрительный нерв** 8](#_Toc59040584)

[**Фоторецепторы** 8](#_Toc59040585)

[**Центральная ямка** 9](#_Toc59040586)

[**Макула** 9](#_Toc59040587)

[**2.2** **Строение и функции палочек и колбочек** 10](#_Toc59040588)

[**2.3 Зрительные пути и зрительная кора** 12](#_Toc59040589)

[**3.** **Обработка зрительного сигнала** 13](#_Toc59040590)

[**3.1** **Обработка информации в сетчатке** 13](#_Toc59040591)

[**Элементы нейронной сети сетчатки и их функции.** 13](#_Toc59040592)

[**Распространение возбуждения в сетчатке.** 13](#_Toc59040593)

[**Электрические реакции нейронов сетчатки.** 14](#_Toc59040594)

[**3.2 Обработка зрительной информации в ЦНС** 15](#_Toc59040595)

[**Сенсорные пути зрительной системы.** 15](#_Toc59040596)

[**Обработка зрительной информации в ЛКТ.** 16](#_Toc59040597)

[**3.3 Обработка информации в зрительных центрах** 17](#_Toc59040598)

[**3.4** **Рефлекс фиксации** 20](#_Toc59040599)

[**3.5** **Движения глаз** 21](#_Toc59040600)

[**Тремор** 21](#_Toc59040601)

[**Дрейф** 21](#_Toc59040602)

[**Микросаккады** 21](#_Toc59040603)

[**Макросаккады** 21](#_Toc59040604)

[**Прослеживающие движения** 22](#_Toc59040605)

[**Вергентные движения** 23](#_Toc59040606)

[**Нистагм** 23](#_Toc59040607)

[**4.** **Распознавание изображений** 24](#_Toc59040608)

[**Глубокое обучение** 24](#_Toc59040609)

[**Нейронная сеть** 24](#_Toc59040610)

[**4.1** **Виды нейронных сетей** 25](#_Toc59040611)

[**Нейронная сеть с прямой связью** 25](#_Toc59040612)

[**4.2** **Когнитрон. Неокогнитрон** 26](#_Toc59040613)

[**Когнитрон** 26](#_Toc59040614)

[**Неокогнитрон** 27](#_Toc59040615)

[**4.3** **Идея использования неокогнитрона** 28](#_Toc59040616)

[**5.** **Заключение** 30](#_Toc59040617)

[**6.** **Литература** 31](#_Toc59040618)

# **Введение**

С повышением наукоёмкости производства в середине ХХ века появилась потребность в машинном зрении – применении технологии создания машин (компьютерного зрения), которые могут производить обнаружение, отслеживание и классификацию объектов. Именно тогда начали появляться первые исследования, касаемые обработки и анализа изображений. В 1958 г. ученый Фрэнк Розенблатт из университета Корнеля создал компьютерную реализацию персептрона, в 1960-ые годы появились первые программные системы обработки изображений, стали развиваться прикладные исследования в области распознавания печатных символов, в 1970-ых годах Лауренс Робертс, аспирант МТИ, выдвинул концепцию машинного построения трехмерных образов объектов на основе анализа их двумерных изображений, в 80-е и 90-е годы появилось новое поколение датчиков двухмерных цифровых информационных полей различной физической природы.

Однако, за последние несколько лет в связи с повышением вычислительных мощностей и быстродействия процессоров, объёмов памяти, разрешающей способности фото-видеотехники, а также с появлением таких технологий, как машинное обучение и искусственный интеллект значимость компьютерного зрения и потребность в применении технологий обработки изображений возрасли.

В настоящее время, компьютерное зрение используется в:

1. Системах управления процессами.
2. Системах видеонаблюдения.
3. Системах организации информации (например, для индексации баз данных изображений).
4. Системах моделирования объектов или окружающей среды (анализ медицинских изображений, топографическое моделирование).
5. Системах взаимодействия (например, устройства ввода для системы человеко-машинного взаимодействия).
6. Системах [дополненной реальности](https://ru.wikipedia.org/wiki/%D0%94%D0%BE%D0%BF%D0%BE%D0%BB%D0%BD%D0%B5%D0%BD%D0%BD%D0%B0%D1%8F_%D1%80%D0%B5%D0%B0%D0%BB%D1%8C%D0%BD%D0%BE%D1%81%D1%82%D1%8C).
7. [Вычислительной фотографи](https://ru.wikipedia.org/wiki/%D0%92%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B8%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D1%84%D0%BE%D1%82%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D1%8F)и, например для мобильных устройств с камерами.

Компьютерное зрение неразрывно связано с биологическим зрением. Технологии машинного зрения созданы по подобию зрительной системы человека. В биологии изучается зрительное восприятие человека и различных животных, в результате чего создаются модели работы таких систем в терминах физиологических процессов. Компьютерное зрение, с другой стороны, изучает и описывает системы, которые выполнены аппаратно или программно.

Следовательно, чтобы успешно решать задачи компьютерного зрения, необходимо понимать, как устроен зрительный анализатор, как работает оптическая система и как происходит обработка зрительного сигнала в мозге. На данные вопросы отвечает нейроинформатика - область научных исследований, лежащая на пересечении нейронаук и информатики. В сферу нейроинформатики входит сбор результатов, полученных в ходе нейробиологических исследований, перевод этих результатов в формат баз данных для их последующего анализа с помощью вычислительных моделей и специализированных компьютерных аналитических программных инструментов, обеспечение совместимости между базами данных, форматами моделей и другими коллекциями данных для облегчения обмена информацией о различных аспектах функционирования и строения нервных систем.

# **Зрительный анализатор**

Рассмотрим анатомическую структуру глаза.

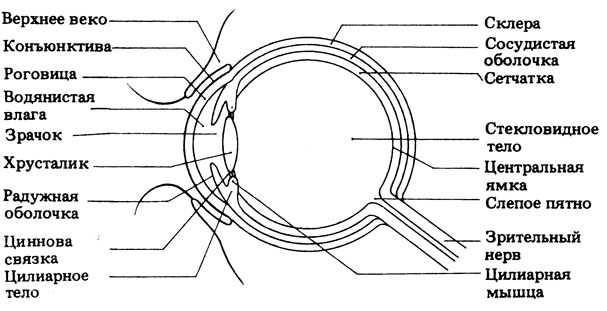


Рис. 1. Схема строения глаза.

**Склера** — это наружная соединительная оболочка, защищающая глаз от мелких повреждений и выполняющая опорную функцию.

**Роговица** — это передняя внешняя прозрачная поверхность склеры, через которую проникает свет. Таким образом, она является главной светопреломляющей структурой.

**Хрусталик** — прозрачное тело, расположенное внутри глазного яблока; является биологической линзой и составляет важную часть светопреломляющего и светопроводящего аппарата глаза.

Пространство между роговицей и хрусталиком заполнено водянистой влагой. Пространство между хрусталиком и сетчаткой заполнено стекловидным телом, которое также является жидкостью, но с большей вязкостью. Давление обеих жидкостей слегка повышено относительно атмосферного, чтобы обеспечить постоянство формы глазного яблока и избежать нежелательного дрожания изображения на сетчатке.

**Радужная оболочка глаза** – часть сосудистой оболочки глаза с отверстием в центре. Практически светонепроницаема.

**Зрачок** – отверстие в радужке, через которое проникает свет.

## **Сетчатка**

Рассмотрим строение и функции сетчатки более подробно.

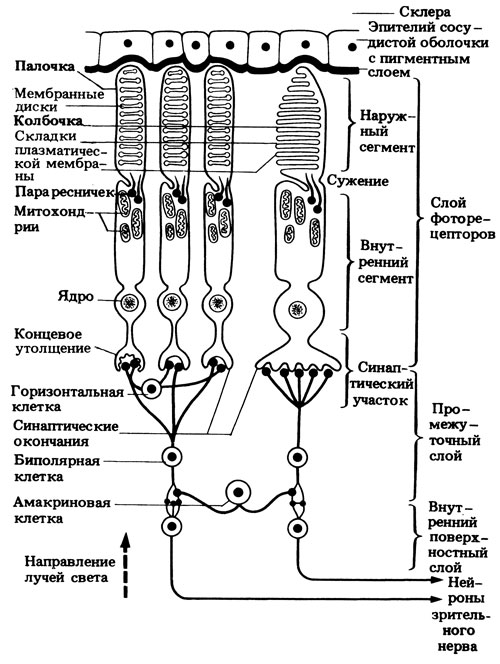


Рис. 2. Схема строения сетчатки.

**Сетчатка** — внутренняя оболочка, содержащая фоторецепторные клетки (палочки для слабой освещённости и колбочки для сильной), а также тела и аксоны нейронов, образующих зрительный нерв.

Сетчатка состоит из трех слоев:

1. Самый наружный содержит **фоторецепторы**, частично погруженные в пигментный слой сосудистой оболочки.
2. Затем идет **промежуточный слой**, содержащий биполярные нейроны, которые связывают фоторецепторы с клетками третьего слоя. В этом же промежуточном слое находятся горизонтальные и амакриновые клетки, обеспечивающие латеральное торможение.
3. Третий слой - **внутренний поверхностный слой** - содержит ганглиозные клетки, дендриты которых соединены синапсами с биполярными клетками, а аксоны образуют зрительный нерв.

**Зрительный нерв** – это вторая пара черепных нервов, по которым зрительные раздражения, воспринятые чувствительными клетками сетчатки, передаются в головной мозг.

Оптический нерв состоит из аксонов (исходящих отростков) ганглиозных клеток — клеток верхнего уровня нейронных процессов в сетчатке.

В зрительный нерв собирается информация со всего пространства, заполненного фоторецепторами, но в каждом глазу есть небольшая область, в кото рой не может быть зрительного возбуждения: эта область носит название слепого пятна.

Оптическое изображение, сформированное глазом, проектируется на сетчатку, которая обеспечивает начальную обработку сигнала и формирует «схему» его передачи в высшие отделы головного мозга.

**Фоторецепторы** – это светочувствительные сенсорные нейроны сетчатки глаза. Фоторецепторы служат для того, чтобы преобразовывать информацию, представленную в оптическом изображении, в химические и электрические сигналы, которые затем могут быть переданы вышестоящим узлам зрительной системы. Эти сигналы обрабатываются сетью нервных клеток сетчатки, а затем передаются в мозг по зрительному нерву.

**Центральная ямка** - наиболее чувствительный участок сетчатки, содержащий только колбочки. В этом участке наиболее точно фокусируются лучи света. Мы двигаем головой и глазами так, чтобы изображение объекта попадало на центральную ямку. Чтобы понять, насколько быстро снижается чёткость изображения по мере отдаления от центральной ямки, можете попытаться прочитать весь текст этого абзаца, зафиксировав взгляд на последней строчке.

**Макула** – центральная часть сетчатки, тонкий слой светочувствительных клеток и нервных волокон. В ней фокусируется свет, когда мы смотрим на какой-то предмет, она несёт ответственность за острое и за цветовое зрения.

## **Строение и функции палочек и колбочек**

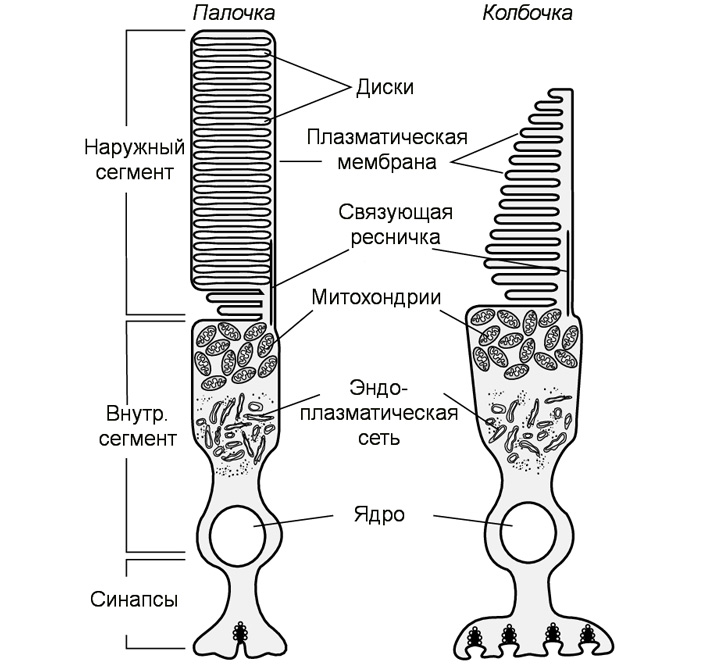


Рис. 3. Схема строения палочек и колбочек.

1. **Наружный сегмент.** Это светочувствительный участок, где световая энергия преобразуется в рецепторный потенциал.
2. **Перетяжка.** Это связь между двумя сегментами, осуществляемая через цитоплазму и пару ресничек, переходящих из одного сегмента в другой.
3. **Внутренний сегмент.** Это область активного метаболизма; она заполнена митохондриями, доставляющими энергию для процессов зрения, и полирибосомами, на которых синтезируются белки, участвующие в образовании мембранных дисков и зрительного пигмента. В этом же участке расположено ядро.
4. **Синаптическая область.** В этом участке клетка образует синапсы с биполярными клетками. Диффузные биполярные клетки могут образовывать синапсы с несколькими палочками. Это явление, называемое синаптической конвергенцией, уменьшает остроту зрения, но повышает светочувствительность глаза. Моносинаптические биполярные клетки связывают одну колбочку с одной ганглиозной клеткой, что обеспечивает большую по сравнению с палочками остроту зрения. Горизонтальные и амакриновые клетки связывают вместе некоторое число палочек или колбочек. Благодаря этим клеткам зрительная информация еще до выхода из сетчатки подвергается определенной переработке; эти клетки, в частности, участвуют в латеральном торможении.

## **2.3 Зрительные пути и зрительная кора**

Нервные импульсы, возникающие в сетчатке, поступают по миллиону волокон зрительного нерва в зрительную кору, расположенную в задней части затылочных долей. В этой зоне спроецированы все мельчайшие участки сетчатки, включающие всего лишь по нескольку палочек и колбочек, и именно здесь зрительные сигналы интерпретируются, что позволяет нам «видеть». Однако то, что мы видим, приобретает смысл только после обмена сигналами с другими участками коры и прежде всего с височными долями, где хранится предшествующая зрительная информация и где она используется для анализа и идентификации текущих зрительных сигналов.

В мозгу человека аксоны от левых половин сетчатки обоих глаз направляются к левой половине зрительной коры, а аксоны от правых половин сетчатки обоих глаз - к правой стороне зрительной коры. Аксоны, идущие от носовых половин обеих сетчаток, пересекаются; место их пересечения называется **зрительным перекрестом** или **хиазмой**. Около 20% волокон зрительного нерва не доходит до зрительной коры, а вступает в средний мозг и участвует в рефлекторной регуляции диаметра зрачка и движений глаз.

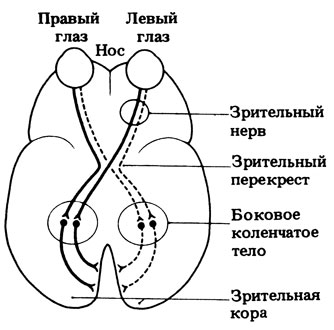


Рис. 4 Строение зрительных путей.

# **Обработка зрительного сигнала**

## **Обработка информации в сетчатке**

### **Элементы нейронной сети сетчатки и их функции.**

Нейронная сеть сетчатки включает 4 типа нервных клеток:

**Ганглиозные клетки**– нейроны, аксоны которых в составе зрительного нерва выходят из глаза и следуют в ЦНС. Функция ганглиозных клеток – проведение возбуждения из сетчатки в ЦНС.

**Биполярные клетки**соединяют рецепторные и ганглиозные клетки. От тела биполярной клетки отходят два разветвленных отростка: один отросток образует синаптические контакты с несколькими фоторецепторными клетками, другой – с несколькими ганглиозными клетками. Функция биполярных клеток – проведение возбуждения от фоторецепторов к ганглиозным клеткам.

**Горизонтальные клетки**соединяют расположенные рядом фоторецепторы. От тела горизонтальной клетки отходит несколько отростков, которые образуют синаптические контакты с фоторецепторами. Основная функция горизонтальных клеток – осуществление латеральных взаимодействий фоторецепторов.

**Амакриновые клетки**расположены подобно горизонтальным, но их образуют контакты не с фоторецепторными, а с ганглиозными клетками.

### **Распространение возбуждения в сетчатке.**

При освещении фоторецептора в нем развивается рецепторный потенциал, который представляет собой гиперполяризацию (изменение мембранного потенциала в отрицательную сторону, процесс обратный деполяризации). Рецепторный потенциал, возникший в фоторецепторной клетке, передается биполярным и горизонтальным клеткам через синаптические контакты с помощью медиатора (биологически активное химическое вещество для передачи нервного импульса от одной клетки к другой).

В биполярной клетке может развиваться как деполяризация, так и гиперполяризация, которая через синаптический контакт распространяется на ганглиозные клетки. Последние являются спонтанно активными, т.е. непрерывно генерируют потенциалы действия с определенной частотой. Гиперполяризация ганглиозных клеток приводит к снижению частоты нервных импульсов, деполяризация – к ее увеличению.

### **Электрические реакции нейронов сетчатки.**

Рецептивное поле биполярной клетки представляет собой совокупность фоторецепторных клеток, с которым она образует синаптические контакты. Под рецептивным полем ганглиозной клетки понимают совокупность фоторецепторных клеток, с которыми данная ганглиозная клетка соединена через биполярные клетки.

Рецептивные поля биполярных и ганглиозных клеток имеют круглую форму. В рецептивном поле можно выделить центральную и периферическую часть. Граница между центральной и периферической частью рецептивного поля является динамичной и может смещаться при изменении уровня освещенности. Реакции нервных клеток сетчатки при освещении фоторецепторов центральной и периферической части их рецептивного поля, как правило, противоположны. При этом существует несколько классов ганглиозных и биполярных клеток, демонстрирующих разные электрические ответы на действие света.

## **3.2 Обработка зрительной информации в ЦНС**

### **Сенсорные пути зрительной системы.**

Миелиновые аксоны ганглиозных клеток сетчатки направляются в головной мозг в составе двух зрительных нервов. Правый и левый зрительные нервы сливаются у основания черепа, образуя зрительный перекрест (хиазму). Здесь нервные волокна, идущие от медиальной половины сетчатки каждого глаза переходят на контрлатеральную сторону(противоположную), а волокна от латеральных половин сетчаток продолжаются ипсилатерально(остаются на той же стороне).

После перекреста аксоны ганглиозных клеток в составе зрительного тракта следуют в латеральные коленчатые тела (ЛКТ), где образуют синаптические контакты с нейронами ЦНС. Аксоны нервных клеток ЛКТ в составе т.н. зрительной лучистости достигают нейронов первичной зрительной коры. Далее по внутрикорковым связям возбуждение распространяется во вторичную зрительную кору и ассоциативные зоны коры.

Сенсорные пути зрительной системы организованы по ретинотопическому принципу (небольшая центральная ямка проецируется на гораздо более обширную область зрительной коры, чем такой же по площади участок с периферии сетчатки) – возбуждение от соседних ганглиозных клеток достигает соседних точек ЛКТ и коры. Поверхность сетчатки как бы проецируется на поверхность ЛКТ и коры.

Большая часть аксонов ганглиозных клеток заканчиваются в ЛКТ, часть же волокон следует в верхние бугры двухолмия, гипоталамус, претектальную область ствола мозга (пограничная область между крышей среднего мозга и промежуточным мозгом, которая участвует в зрительном анализе движущихся объектов), ядро зрительного тракта.

•  Связь между сетчаткой и верхними буграми четверохолмия служит для регуляции движений глаз.

•  Проекция сетчатки в гипоталамус служит для сопряжения эндогенных циркадных ритмов с суточными колебаниями уровня освещенности.

•  Связь между сетчаткой и претектальной областью ствола исключительно важна для регуляции просвета зрачка и аккомодации.

•  Нейроны ядер зрительного тракта, которые также получают синаптические входы от ганглиозных клеток, связаны с вестибулярными ядрами ствола мозга. Эта проекция позволяет оценивать положение тела в пространстве на основании зрительных сигналов, а также служит для осуществления сложных глазодвигательных реакций (нистагм).

### **Обработка зрительной информации в ЛКТ.**

•  Нейроны ЛКТ имеют рецептивные поля округлой формы. Электрические реакции этих клеток аналогичны таковым ганглиозных клеток.

•  В ЛКТ существуют нейроны, которые возбуждаются при наличии в их рецептивном поле границы свет/темнота (контрастные нейроны) или при передвижении этой границы в пределах рецептивного поля (детекторы движения).

## **3.3 Обработка информации в зрительных цен­трах**

Основной анализ зрительной информации соверша­ется нейронами коры, среди которых выделяют простые, сложные, сверхсложные и гностические. Все эти нейроны объединены в вертикальные колонки (глазодоминантные, ориентационные).

Прежде всего, потенциалы действия нейронов наруж­ного коленчатого тела поступают в затылочную часть по­лушарий большого мозга, где расположена первичная проекционная область зрительной зоны коры (стриарная кора, или поле 17). Афферентные волокна (чувствительные волокна, проводящие импульсы от периферии к ЦНС) из наружного коленчатого тела оканчиваются в слое IV и в глубине слоя III 17-го поля, которое является центральным по­лем зрительной коры, а 18-е и 19-е поля - перифери­ческими. Ассоциативные волокна из поля 17 направляют­ся в 18-е и 19-е поля. Между сетчаткой и полем 17 су­ществует упорядоченное топографическое соответствие, благодаря чему карта сетчатки, а, следовательно, и поле зрения проецируется на кору. Центральная ямка, где ос­трота зрения максимальна, занимает большую часть кор­кового представительства. Таким образом, единственную точную ретинооптическую карту содержит только поле 17. Смежным с ним полям 18 и 19 приписывают нето­пографические ассоциативные функции.

В первичной проекционной зоне происходит анализ информации, поступающей одновременно из правого и ле­вого глаз. Как и в других зонах коры, в этой области анализ проводится с участием колонок. Имеются глазодо­минантные колонки, анализирующие информацию, идущую либо из правого глаза, либо из левого. Эти колонки со­седствуют друг с другом, поэтому, вероятно, между ними происходит обмен информацией, и это позволяет видеть двумя глазами один предмет (бинокулярное зрение).

Среди нейронов колонок коры различают «простые» нейроны, задача которых - выявить контраст, наличие движущегося стимула, т. е. точно такие же задачи, как у нейронов сетчатки (но для нейронов коры рецептивные поля имеют более обобщенный характер). Простые нейроны зрительной коры в целом выполняет свои специфи­ческие функции, получая сигналы от определенного учас­тка сетчатки и откликаясь максимальной активностью на определенную форму и ориентацию стимула.

Кроме того, колонки зрительной коры содержат «сложные» и «сверхсложные» нейроны, которые возбуж­даются при наличии определенных условий, например, при движении светового сигнала слева направо или снизу вверх (дирекциональные детекторы), либо возбуждаются на определенный цвет, часть нейронов лучше всего отвеча­ет на относительную удаленность объекта от глаз. Важно подчеркнуть, что информация о разных признаках зри­тельных объектов (форма, цвет, движение) обрабатывается параллельно в разных частях зрительной зоны коры боль­шого мозга.

Таким образом, в первичной проекционной зоне про­исходит более специализированная и сложная, чем в сет­чатке и в латеральном коленчатом теле, переработка ин­формации. Нейроны зрительной зоны коры имеют не круглые, а вытянутые (по горизонтали, вертикали или в одном из косых направлений) рецептивные поля неболь­шого размера. Благодаря этому нейроны коры способны выделять из цельного изображения отдельные фрагменты линий с той или иной ориентацией и расположением (де­текторы ориентации) и избирательно на них реагировать.

Вся информация от нейронов первичного (поле 17) и вторичных (поля 18 и 19) проекционных полей коры пе­редается в передние и задние ассоциативные зоны коры (лобные и теменно-височные области коры), где с учас­тием более «обученных» (гностических) нейронов проис­ходит окончательное формирование образа. В частности, в ассоциативных полях содержатся нейроны, обученные узнавать все буквы алфавита, слова, лица и т. д. Когда соответствующий нейрон «узнает» предназначенное ему для узнавания, он возбуждается, и это является физио­логическим механизмом восприятия.

## **Рефлекс фиксации**

Если в поле зрения попадает какой-нибудь предмет, на нём непроизвольно фиксируется взгляд. При движении предмета глаза следуют за ним, при этом изображение предмета фиксируется в точке наилучшего видения на сетчатке, то есть в зоне ямок жёлтых пятен. Когда мы произвольно рассматриваем интересующий нас предмет, взгляд автоматически задерживается на нём, даже если мы сами или предмет движется. Таким образом, произвольные движения глаз основаны на непроизвольных рефлекторных движениях. Этот рефлекс — фиксирование изображения интересующего объекта на сетчатке в зоне наиболее чёткого видения — называется рефлексом фиксации.

Афферентный путь (чувствительные волокна) этого рефлекса идёт от сетчатки по зрительным путям к зрительной коре (поле 17 — затылочная доля). Оттуда импульсы передаются в зоны 18 и 19 (затылочная доля). Эфферентные (двигательные) волокна, вероятно, возникают именно в этих зонах, затем временно присоединяются к волокнам зрительной лучистости, следуя к контралатеральным глазодвигательным центрам моста и среднего мозга. Отсюда волокна идут к соответствующим ядрам двигательных нервов глаза. Вероятно, некоторые эфферентные волокна идут прямо к глазодвигательным центрам.

Сегменты верхних бугорков четверохолмия также могут быть ответственны за движения глазных яблок в определённых направлениях. Центры, ответственные за движения кверху, находятся в передних отделах верхних бугорков. Импульсы, возникающие в полюсах затылочных долей, также передаются в контралатеральные глазодвигательные центры моста и вызывают содружественные боковые движения глазных яблок.

## **Движения глаз**

Известны восемь основных видов движений глаз: тремор, дрейф, микро- и макросаккады, прослеживающие вергентные, торзионные движения и нистагм.

**Тремор** – мелкие, частые колебания глаз. Средняя амплитуда – 20–40", частота – до 250–270 Гц. В результате тремора ось глаза описывает эллипсоподобные фигуры. Тремор – естественный двигательный фон окуломоторной активности, неподдающийся произвольному контролю. Возникает при фиксации взора на объекте для его дальнейшего рассмотрения.

**Дрейф** – медленное, плавное перемещение глаза, прерываемое микросаккадами. Скорость дрейфа меняется от 0 до 40°/с, длительность – от 30 до 5000 мс. Считается, что дрейф создает наиболее благоприятные условия для приема и переработки оптической информации.

**Микросаккады** – быстрые резкие движения продолжительностью 10–20 мс. Диапазон амплитуд – 2–50', скорость от 3 до 12°/с. Микросаккады плохо поддаются произвольному контролю, появляясь во время фиксации объектов. Периодические «сплывы» глаз, компенсируемые микросаккадами, образуют самостоятельную двигательную единицу – физиологический нистагм.

Тремор, дрейф и малоаплитудные саккады относят к категории микродвижений глаз, противопоставляя их макродвижениям: крупноамплитудным саккадам, прослеживающим (следящим) и вергентным движениям. Если микродвижения связаны преимущественно с сохранением, то макродвижения – с изменением положения глаз в орбитах.

**Макросаккады** – резкие изменения позиции глаза, отличающиеся высокой скоростью и точностью. Амплитуда саккад варьирует в широких пределах от 40–50' до 50–60°, но в естественных условиях восприятия редко превышает 20°. Продолжительность, скорость и ускорение саккад зависят от их амплитуды. Скорость саккады плавно достигает максимума (примерно в середине пути) и затем плавно убывает до 0. Максимальная скорость двадцатиградусного скачка – 450°/с, его продолжительность – 70 мс. Средняя частота саккадических движений – 2–3 Гц. Как правило они совершаются по кратчайшей прямой между смежными точками фиксации, но в принципе их траектория может иметь синусоидальную, крючкообразную и другие неправильные формы. Саккады возникают при смене точек фиксации, например, во время рассматривания картины, поиска заданного объекта, пересчета элементов и др., и обычно носят произвольный характер (имеется в виду произвольность выбора наблюдателем нового объекта фиксации; произвольно изменить продолжительность, скорость или ускорение макросаккады невозможно). В момент скачка складываются наименее благоприятные условия для получения оптической информации (эффект «саккадического подавления»).

**Прослеживающие движения** – плавные перемещения глаз, возникающие при движении объекта в поле зрения. Они обеспечивают сохранение его изображения в зоне наилучшего видения. Прослеживающие движения глаз появляются непроизвольно через 150–200 мс после начала движения объекта и продолжаются в течение 300 мс после его остановки или исчезновения. Основной диапазон скоростей – 5–90°/с. Амплитуда движений ограничивается пределами моторного поля глаза (около ±60° по горизонтали и ±40° по вертикали). За небольшим исключением вызвать прослеживающие движения произвольно (например, по представлению движущегося объекта) невозможно. Прослеживающие движения глаз существенно расширяют диапазон скоростей движения объектов, в котором сохраняется возможность эффективного восприятия. Другим источником плавных движений глаз являются повороты головы. Однако в этом случае параметры движений глаз могут вообще не зависеть от свойств оптической стимуляции (они сохраняются и в темноте).

**Вергентные движения** – сведение (конвергенция) или разведение (дивергенция) оптических осей глаз. Они включены в процесс бинокулярного зрения, обеспечивая необходимое соответствие проекций объекта на сетчатках обоих глаз: при фиксации зрительные оси пересекаются на наблюдаемом объекте. Стимулом вергентных движений является диспаратность и диплопия (раздвоение) изображений нового объекта фиксации, вследствие раздражения несимметричных областей сетчаток. Движения возникают через 200 мс после появления объекта и продолжаются несколько сот миллисекунд. Они носят преимущественно плавный характер с максимальной скоростью в несколько десятков угловых градусов в секунду. В момент конвергенции (дивергенции) зрительная способность глаз сохраняется.

Торзионные, или **ротационные движения** – вращательные перемещения глаз относительно оптической оси. Они содержат как саккадический, так и плавный компонент; амплитуда движений ограничена 10°. Основное назначение – частичная компенсация наклонов головы относительно гравитационной вертикали.

**Нистагм** – устойчивая окуломоторная структура, включающая чередование саккад и плавных прослеживающих движений. Амплитуда, частота и форма нистагма широко варьируют в пределах параметров базовых видов движений и зависят от его природы; нистагм может иметь оптическое (например, оптокинетический нистагм), эхоическое (вестибулярный нистагм), центральное (виды врожденного нистагма) происхождение. Функция нистагма – компенсация нарушений зрительной и вестибулярной систем или сохранение эффективности восприятия элементов движущейся среды, имеющей регулярную структуру.

# **Распознавание изображений**

Человек легко может распознать знакомый ему объект, даже если он искажён, повёрнут на 90 градусов, или даже если у этого изображения отсутствуют какие-то части. Например, нетрудно догадаться, что на всех шести картинках (Рис. 5) представлена буква «А», пусть и искажённая.

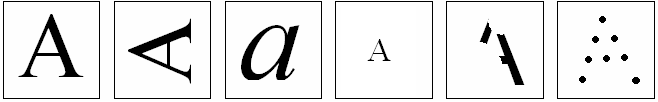


Рис. 5 Искажённая буква «А».

Тем не менее, для компьютера данная задача может оказаться очень сложной или даже неразрешимой. Тогда на помощь приходит машинное зрение. Распознавание изображений – одна из классических задач данной технологии.

**Глубокое обучение** является частью методов машинного обучения, основанных на искусственных нейронных сетях.

**Нейронная сеть** (также искусственная нейронная сеть, ИНС) — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.

## **Виды нейронных сетей**

Существует достаточно большое количество разновидных нейросетей. Но мне бы хотелось заострить внимание на этих двух:

**Нейронная сеть с прямой связью** — искусственная нейронная сеть, в которой соединения между узлами не образуют цикл. В этой сети информация перемещается только в одном направлении вперед от входных узлов, через скрытые узлы (если они есть) и к выходным узлам. В сети нет циклов или петель обратных связей.

Но данная нейросеть является неустойчивой к небольшим сдвигам, поворотам и изменению масштаба изображения.

Решение этой проблемы было найдено ученым Яном ЛеКуном, вдохновленным работами нобелевских лауреатов в области медицины Торстена Нильса Визеля и Дэвида Хьюбела. Эти ученые исследовали зрительную кору головного мозга кошки и обнаружили, что существуют так называемые простые клетки, которые особо сильно реагируют на прямые линии под разными углами и сложные клетки, которые реагируют на движение линий в одном направлении. Ян ЛеКун предложил использовать так называемые **Сверточные нейронные сети.** Идея свёрточных нейронных сетей заключается в чередовании свёрточных слоёв (включает в себя для каждого канала свой фильтр, ядро свёртки которого обрабатывает предыдущий слой по фрагментам (суммируя результаты поэлементного произведения для каждого фрагмента) и субдискретизирующих слоёв (слоёв подвыборки, где происходит процесс сжатия (уменьшения размеров) изображения путём сложения значений блоков пикселей).

## **Когнитрон. Неокогнитрон**

Данные нейронные сети одновременно является как моделью процессов восприятия на микроуровне, так и вычислительными системыми, применяющейся для технических задач распознавания образов.

**Когнитрон** — искусственная нейронная сеть на основе принципа самоорганизации (процесс упорядочения элементов одного уровня в системе за счёт внутренних факторов, без внешнего специфического воздействия). Своей архитектурой когнитрон похож на строение зрительной коры. В зрительной коре были обнаружены узлы, реагирующие на такие элементы, как линии и углы определенной ориентации. На более высоких уровнях узлы реагируют на более сложные и абстрактные образы такие, как окружности, треугольники и прямоугольники. На ещё более высоких уровнях степень абстракции возрастает до тех пор, пока не определятся узлы, реагирующие на лица и сложные формы. В общем случае узлы на более высоких уровнях получают вход от группы низкоуровневых узлов и, следовательно, реагируют на более широкую область визуального поля. Реакции узлов более высокого уровня менее зависят от позиции и более устойчивы к искажениям.

Обучается конкурентным обучением (без учителя). Каждый слой мозга реализует различные уровни обобщения; входной слой чувствителен к простым образам, таким, как линии, и их ориентации в определенных областях визуальной области, в то время как реакция других слоев является более сложной, абстрактной и независимой от позиции образа. Аналогичные функции реализованы в когнитроне путём моделирования организации зрительной коры.

Когнитрон состоит из иерархически связанных слоев нейронов двух типов — тормозящих и возбуждающих. Состояние возбуждения каждого нейрона определяется соотношением его тормозящих и возбуждающих входов. Синаптические связи идут от нейронов одного слоя (далее слоя 1) к следующему (слою 2). Относительно данной синаптической связи соответствующий нейрон слоя 1 является пресинаптическим (передающий нейрный импульс), а нейрон второго слоя — постсинаптическим (принимающий этот импульс). Постсинаптические нейроны связаны не со всеми нейронами 1-го слоя, а лишь с теми, которые принадлежат их локальной области связей. Области связей близких друг к другу постсинаптических нейронов перекрываются, поэтому активность данного пресинаптического нейрона будет сказываться на все более расширяющейся области постсинаптических нейронов следующих слоев иерархии.

Когнитрон конструируется в виде слоев нейронов, соединенных синапсами. Пресинаптический нейрон в одном слое связан с постсинаптическим нейроном в следующем слое. Имеются два типа нейронов: возбуждающие узлы, которые стремятся вызвать возбуждение постсинаптического узла, и тормозящие узлы, которые тормозят это возбуждение. Возбуждение нейрона определяется взвешенной суммой его возбуждающих и тормозящих входов, однако в действительности механизм является более сложным, чем простое суммирование.

**Неокогнитрон** является дальнейшим развитием идеи когнитрона и более точно отражает строение зрительной системы, позволяет распознавать образы независимо от их преобразований, вращений, искажений и изменений масштаба. Неокогнитрон может как самообучаться, так и обучаться с учителем. Неокогнитрон получает на входе двумерные образы, аналогичные изображениям на сетчатке глаза, и обрабатывает их в последующих слоях аналогично тому, как это было обнаружено в зрительной коре человека.

## **Идея использования неокогнитрона**

Одна из проблем машинного обучения – простота взлома. Если воздействовать на процесс обучения нейросети, например, наложить на изображение панды специальный «шум», то она распознает в панде гиббона, причём с очень большой уверенностью (см рис). Тем не менее, даже готовую математическую модель можно взломать, если знать её архитектуру и принцип работы.

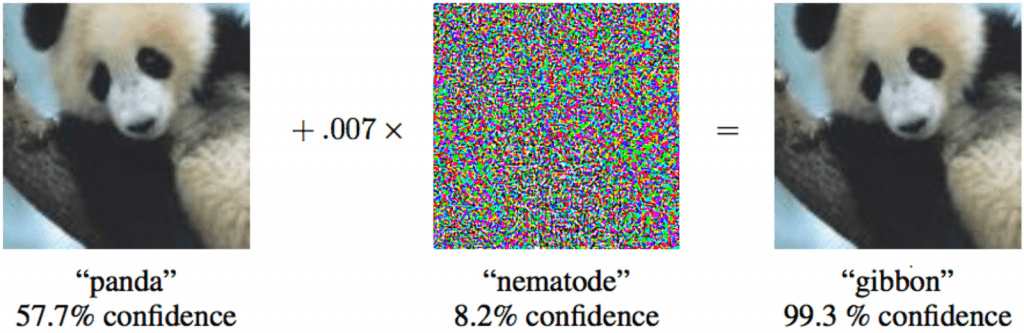


Рис. 6 Нейросеть распознала панду как гиббона.

Неокогнитрон своим строением похож на зрительную кору (см п.4.2), таким образом, он имеет высокую точность вычислений и устойчивость к искажению изображений. Поэтому он не так восприимчив к таким «обманам», как другие виды нейросетей. Также данная система имеет способность к самовосстановлению при обучении без учителя: если какой-то узел выйдет из строя, то буден найден другой, реагирующий сильнее, и будет обучен распознаванию входного образа, перекрывая любые действия отказавшего узла.

Данный вид нейросети может найти применение в любой области, он является достаточно универсальным. Его применяют как для решения технических задач распознавания изображений, так и для реализации каких-либо сложных вычислений.

Конкретный пример использования: создание программы для анализа медицинских изображений, полученных методом лучевой диагностики (рентгенологическим, магнитно-резонансным, радионуклидным или ультразвуковым). При обработке и анализе таких изображений выделяют следующие этапы: фильтрация; предварительная обработка; сегментация; распознавание; диагностика. Человеку иногда бывает сложно выявить патологию по изображению, он может упустить какие-то детали или же вовсе не увидеть того, что может заметить машина. Поэтому точность распознавания неокогнитрона может быть полезна для решения задачи анализа медицинских изображений.

К сожалению, использовать неокогнитрон для решения сегодняшних проблем распознавания изображений неоптимально, так как реализация данной нейросети требует существенных вычислительных ресурсов. Но это не повод отказываться от дальнейшего изучения данного вида нейросети, потому что тенденция уменьшения стоимости вычислений будет сохраняться как минимум ещё 10 лет.

Возможно, в ближайшем будущем людям удастся упростить вычисления без потерь точности или же создать модель зрительного анализатора более близкую к биологической. Следовательно, искусственная нейронная сеть, построенная на основе зрительной коры, имеет шансы на развитие и является перспективной за счёт высокой точности.

# **Заключение**

В ходе проведённого мной исследования, была рассмотрена модель зрительной системы человека, процесс обработки зрительного сигнала в мозге, а также некоторые виды искусственных нейронных сетей (в том числе, когнитрон и неокогнитрон), архитектура которых основана на строении зрительной коры человека, для решения задач распознавания изображений. Было установлено, что данные нейросети имеют высокую точность вычислений и устойчивость к различным искажениям изображений (добавлению шума, фильтрам, изменению масштаба и т. д.). Но они не лишены недостатков – а именно большой вычислительной сложности. Следовательно, нейросети – прототипы зрительной системы человека эффективны при решении задач распознавания изображений, где требуется высокая точность, например, в медицине. Тем не менее, вопрос о снижении сложности вычислений, при обучении данных нейросетей остаётся открытым.

# **Литература**

1. Компьютерное зрение: <https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80%D0%BD%D0%BE%D0%B5_%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5>
2. Машинное зрение: <https://ru.wikipedia.org/wiki/%D0%9C%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5_%D0%B7%D1%80%D0%B5%D0%BD%D0%B8%D0%B5>
3. Системы машинного зрения:

<https://robotics.ua/shows/modernity/5844-sistemy_mashinnogo_zreniya_istoriya_primery_plany>

1. Нейроинформатика: <https://ru.wikipedia.org/wiki/%D0%9D%D0%B5%D0%B9%D1%80%D0%BE%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0>
2. Строение глаза: <https://ru.wikipedia.org/wiki/%D0%93%D0%BB%D0%B0%D0%B7>
3. Строение глаза: <http://biologylib.ru/books/item/f00/s00/z0000018/st265.shtml>
4. Обработка зрительной информации: <https://studopedia.su/20_94270_obrabotka-zritelnoy-informatsii.html>
5. Физиология зрения:

<http://www.bio.bsu.by/phha/19/19_text.html#02>

1. Поля Бродмана: <https://ru.wikipedia.org/wiki/%D0%A6%D0%B8%D1%82%D0%BE%D0%B0%D1%80%D1%85%D0%B8%D1%82%D0%B5%D0%BA%D1%82%D0%BE%D0%BD%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B5_%D0%BF%D0%BE%D0%BB%D1%8F_%D0%91%D1%80%D0%BE%D0%B4%D0%BC%D0%B0%D0%BD%D0%B0#:~:text=%D0%9F%D0%BE%D0%BB%D1%8F%20%D0%91%D1%80%D0%BE%D0%B4%D0%BC%D0%B0%D0%BD%D0%B0%20%E2%80%94%20%D0%BE%D1%82%D0%B4%D0%B5%D0%BB%D1%8B%20%D0%BA%D0%BE%D1%80%D1%8B%20%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D1%85,%D0%BA%D0%BE%D1%80%D1%8B%20%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B8%D1%85%20%D0%BF%D0%BE%D0%BB%D1%83%D1%88%D0%B0%D1%80%D0%B8%D0%B9%20%D0%B3%D0%BE%D0%BB%D0%BE%D0%B2%D0%BD%D0%BE%D0%B3%D0%BE%20%D0%BC%D0%BE%D0%B7%D0%B3%D0%B0>.
2. Неокогнитрон: <https://ru.wikipedia.org/wiki/%D0%9D%D0%B5%D0%BE%D0%BA%D0%BE%D0%B3%D0%BD%D0%B8%D1%82%D1%80%D0%BE%D0%BD>
3. Когнитрон: <https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%B3%D0%BD%D0%B8%D1%82%D1%80%D0%BE%D0%BD>
4. Нейронная сеть с прямой связью: <https://ru.wikipedia.org/wiki/%D0%9D%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C_%D1%81_%D0%BF%D1%80%D1%8F%D0%BC%D0%BE%D0%B9_%D1%81%D0%B2%D1%8F%D0%B7%D1%8C%D1%8E>
5. Сверточные нейросети: <https://ru.wikipedia.org/wiki/%D0%A1%D0%B2%D1%91%D1%80%D1%82%D0%BE%D1%87%D0%BD%D0%B0%D1%8F_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C>
6. Проблемы машинного обучения: <https://www.kaspersky.ru/blog/machine-learning-ten-challenges/21193/>
7. Неокогнитрон и когнитрон: <https://intuit.ru/studies/courses/88/88/lecture/20553?page=2>
8. Обработка зрительного сигнала: <https://studfile.net/preview/5836349/page:9/>
9. Компьютерная модель зрительной системы:

<http://www.jip.ru/2005/414-425.pdf>

1. Нейросети:

<https://habr.com/ru/post/74326/>

1. Машинное зрение:

<https://habr.com/ru/post/350918/>

1. Движения глаз в процессе фиксации (А.Л. Ярбус 1966г.):

<http://asenic.ru/ocrlab/yarbus_fin.htm>

1. Движения глаз:

<https://lib.ipran.ru/upload/papers/23345071.pdf>

1. Рефлекс фиксации: https://ru.wikipedia.org/wiki/%D0%94%D0%B2%D0%B8%D0%B6%D0%B5%D0%BD%D0%B8%D0%B5\_%D0%B3%D0%BB%D0%B0%D0%B7