

К ВОПРОСУ ВОСПРИЯТИЯ И РАСПОЗНАВАНИЯ ОБРАЗОВ В СИСТЕМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Анотація. У статті в рамках біонічного підходу розглянуто питання щодо сприйняття й розпізнавання образів у системах штучного інтелекту. Представлено опис пристрою сенсорної зорової системи сприйняття і попередньої обробки інформації – очі людини. Висунуто робочі гіпотези «Про приведення зображення, що розпізнається, до одного розміру в області фовеа» і «Про механізм розпізнавання образів у вищих шарах неокортекса мозку людини». Представлено опис створеної на базі робочих гіпотез апаратно-нейронної моделі сенсорного органа зорової системи людини. Описана нейронна мережа виділення контуру зображення, що розпізнається.

Ключові слова: біонічний підхід, сприйняття інформації, розпізнавання образів, гіпотеза, модель зорової системи людини.

Аннотация. В статье в рамках бионического подхода рассмотрены вопросы восприятия и распознавания образов в системах искусственного интеллекта. Представлено описание устройства и функционирования сенсорной зрительной системы восприятия и предварительной обработки информации – глаза человека. Выдвинуты рабочие гипотезы «О приведении распознаваемых изображений к одному размеру в области фовеа» и «О механизме распознавания образов в высших слоях неокортекса мозга человека». Представлено описание созданной на базе рабочих гипотез аппаратно-нейронной модели сенсорного органа зрительной системы человека. Описана простая нейронная сеть для выделения контура распознаваемого изображения.

Ключевые слова: бионический подход, восприятие информации, распознавание образов, гипотеза, модель зрительной системы человека.

Abstract. The question of pattern perception and identification in the artificial intelligence systems was regarded in this article in the context of bionic approach. A description of a sensor visual system perception and information preprocessing device – human eyes was represented. The working hypotheses were put forward “About working the recognizable image out to the same size in the field of fovea” and “About the mechanism of pattern recognition in the higher layers of the neocortex of the human brain”. The description created on the basis of working hypotheses of hardware and neural models of visual sensory organ system was represented. The neural network of recognizable image contour detection is described.

Keywords: bionic approach, information perception, pattern recognition, hypothesis, human system visual model.

1. Введение

Создание систем распознавания образов с элементами искусственного интеллекта является

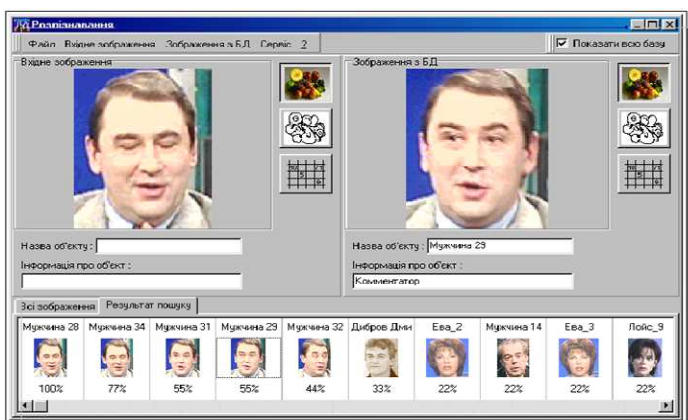


Рис. 1. Интерфейс системы распознавания

сложной теоретической и технической проблемой. Необходимость в таком распознавании возникает в самых разных областях.

Традиционно задачи распознавания образов включают в круг задач искусственного интеллекта, где выделяют два основных направления:

первое – развитие теории и методов построения программных комплексов, предназначенных для решения отдельных задач распозна-

вания в прикладных целях;

второе – изучение механизмов распознавания, которыми обладают живые существа, с целью их воспроизведения в системах распознавания.

Являясь сторонником второго направления, автор, совместно со студентами Киевского национального университета имени Тараса Шевченко, порядка одиннадцати лет назад, тщательно изучив существующую на то время литературу по нейрофизиологии, создали программную систему распознавания лиц, которая была соединена с TV, видеокамерой, и в режиме реального времени успешно распознавала изображения лиц телеведущих, изображения с видеокамеры и фотографии лиц из базы фотоснимков (The Yale Face Database и The PICS images database). На рис. 1 показан интерфейс системы, на котором виден результат процесса распознавания телеведущего того времени. Слева показано изображение телеведущего, полученное из TV в реальном времени, справа его изображение из базы изображений системы распознавания лиц, запомненное при обучении. В реализации системы применен бионический подход к решению задачи восприятия и распознавания образов.

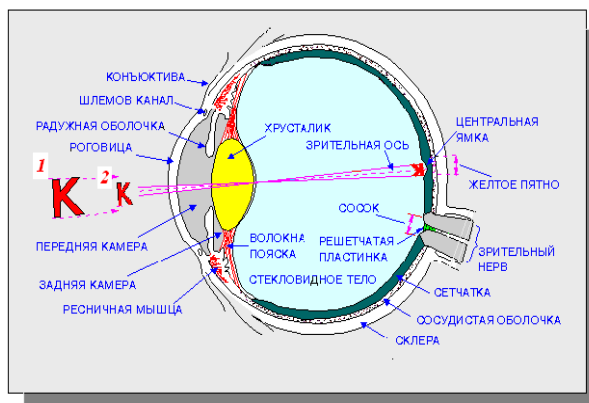
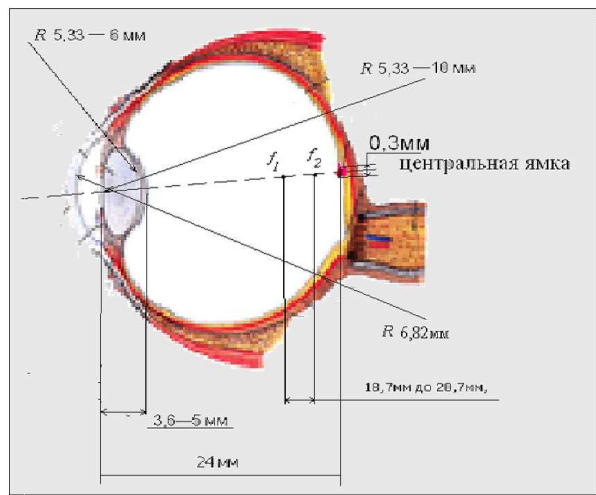
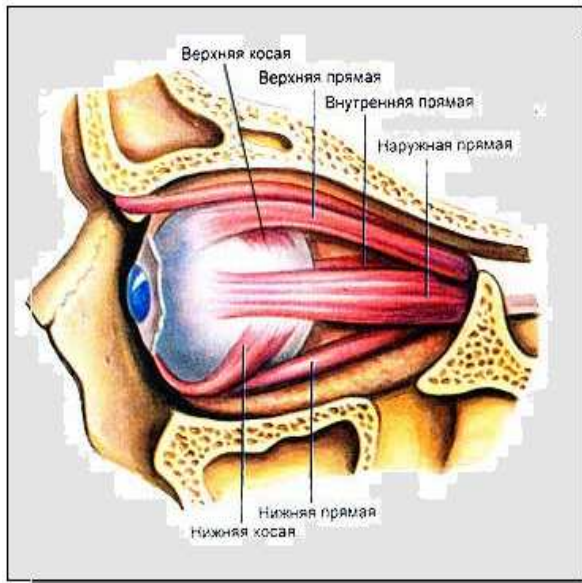


Рис. 2. Схема глаза человека

2. Бионический подход к решению задачи восприятия и распознавания образов

При бионическом подходе к решению задачи восприятия и распознавания образов взоры исследователей обращаются к зрительной системе человека. Тем более что 80–90 % воспринимаемой информации поступает через глаза.

2.1. Сенсорный орган зрительной системы человека

Глаз – сенсорный орган зрительной системы человека состоит из глазного яблока и зрительного нерва (рис. 2).

Глазное яблоко имеет диаметр около 24 мм и форму почти правильного шара, но с чуть выпуклой передней частью.

Вокруг каждого глазного яблока находится шесть мышц. Глаза двигаются наружной и внутренней прямыми мышцами: влево, вправо; нижней и верхней прямыми мышцами – вниз, вверх; косые мышцы вращают глаз. Но на этом работа глазодвигательных мышц не заканчивается. 30-

летние исследования У. Бейтса доказали, что прямые мышцы глаза могут укорачивать глазное яблоко вдоль оптической оси глаза, таким образом приближая хрусталик к сетчатке, а косые – могут сжимать глаз и отодвигать хрусталик от сетчатки.

Наружная оболочка глазного яблока состоит из плотной части – склеры, белкового вещества, придающего форму глазу и защищающего внутреннюю часть глаза от воздействий внешней среды. Склера участвует также в обмене веществ.

Передняя часть склеры переходит в прозрачный круг – роговицу. Роговица имеет радиус кривизны 6,82 мм и выполняет роль собирающей линзы.

Между роговицей и хрусталиком расположена радужная оболочка. Радужная оболочка снабжена мышцами, нервными волокнами, кровеносными сосудами. Радужная оболочка глаза регулирует силу светового потока таким образом, что на рецепторы глаза поступает почти постоянный световой поток (при сильном освещении радужная оболочка расширяется, суживая зрачок, при слабом освещении она сужается, расширяя зрачок). Пространство, образованное между роговицей и радужкой, называется передней камерой. Она заполнена внутриглазной жидкостью.

Через хрусталик изображение проецируется на сетчатку глаза (ретину). Благодаря аккомодации – эластичности хрусталика и действию глазной мышцы, хрусталик приобретает форму, обеспечивающую резкую проекцию изображения на светочувствительные преобразователи (рецепторы), расположенные в сетчатке глаза.

Процесс аккомодации состоит в изменении оптической силы (рефракции) глаза. Это достигается изменением радиуса кривизны хрусталика в результате действия кольцеобразной мышцы. При расслабленной мышце связки натягивают мешочек хрусталика и кривизна его поверхностей становится наименьшей. В этом случае на сетчатке получается резкое изображение удаленных предметов. Фокусное расстояние глаза взрослого человека изменяется в пределах от 18,7 мм до 20,7 мм, что обеспечивает фокусировку как на дальних, так и на ближних объектах. При максимальном сжатии мускульного кольца глаз отчетливо видит наиболее близкие предметы. Формирование изображения в основном осуществляется роговицей вместе с хрусталиком, которые в комбинации имеют фокусное расстояние около 20 мм.

Сетчатка состоит из множества отдельных элементов – рецепторов, каждый из которых реагирует на световой поток независимо друг от друга. Глаз человека содержит светочувствительные элементы двух типов: колбочки и палочки. Распределение палочек и колбочек в сетчатке неравномерно: палочек больше на периферии, а колбочек – в центре.

В центре сетчатки имеется участок, содержащий только колбочки. Этот участок называется центральной ямкой (область фовеа). Разрешающая способность глаза здесь максимальна, причем колбочки центральной ямки непосредственно связаны с высшими нервными центрами, в то время как большинство рецепторов сетчатки не имеют «прямой» связи с мозгом. Можно полагать, что основная часть работы по извлечению зрительной информации выполняется той частью рецепторов, которые лежат в середине сетчатки в области центральной ямки.

Зрительный центр головного мозга находится в затылочной части головы, в основании черепа. Основная часть нервных путей пересекается в передней части основания черепа. Мозг принимает информацию из правого и левого глаза. Затем она объединяется в единый образ. Как раз для пространственного зрения важно взаимодействие обоих глаз, чтобы глаза были направлены в одну точку. Так возникает изображение, в противном же случае увиденное раздваивается.

Ранее была выдвинута гипотеза «О приведении распознаваемых изображений к одному размеру в области фовеа», реализация которой, на мой взгляд, подтверждается исследованиями офтальмологов.

Рабочая гипотеза 1. Гипотеза о приведении распознаваемых изображений к одному размеру в области фовеа. Учитывая, что в системе «глаз человека» имеется только одна область (центральная ямка, фовеа), где разрешающая способность глаза максимальна, причем колбочки центральной ямки непосредственно связаны с высшими нервными центрами, можно предположить, что распознаваемый объект (объект, на котором сконцентрировано внимание, обращен взгляд), например, «К1» или «К2» (рис. 2, 3), сканируется саккадами – движениями глаз, осуществляя систематический отбор информации о форме, положении и размере объекта, проецируется в область фовеа с учетом этих параметров, затем в высших отделах мозга происходит анализ, синтез и сравнение с запомненными ранее

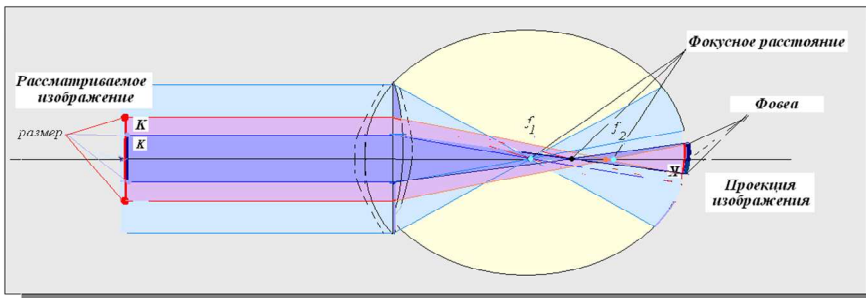


Рис. 3. Схема проекции изображения в область фовеа

объектами по уровню возбуждения нейронов, отражающих признаки и свойства этих объектов. Если размер объекта большой и его проекция выходит за пределы области фовеа, то некоторая его часть, выходящая за пределы фовеа, будет видна не четко. При нормальном функционировании глаза этого не происходит,

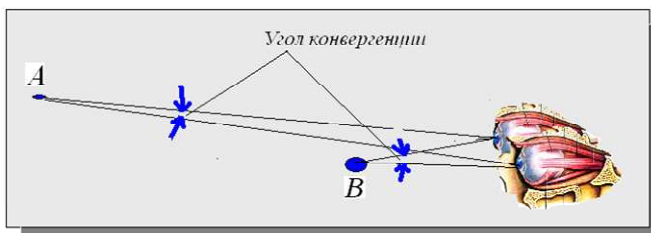


Рис. 4. Конвергенция глаз

значит, проекция должна полностью проецироваться в область фовеа за счет изменения фокусного расстояния. А размеры объектов могут определяться не размерами их проекций в области фовеа, а уровнем возбуждения соответствующих командных нейронов глазных мышц, управляющих формой хрусталика¹, или, по другим источникам, формой глазного яблока², что приводит к изменению фокусного расстояния линзы – хрусталика в первом случае, а во втором – к изменению фокусного расстояния системы глаз. Кроме того, оценивание расстояния обоими глазами осуществляется за счет конвергенции глаз (состояния командных нейронов внутренних прямых глазных мышц). Так, при смотре обоими глазами в одну и ту же точку А, как показано на рис. 4, необходимо некоторое мускульное усилие внутренних прямых глазных мышц для того, чтобы свести оба глаза внутрь. Глаза расположены на некотором расстоянии друг от друга, и мы на опыте учимся оценивать расстояние до точки А по мускульному усилию, необходимому для сведения (конвергенции) глаз. Для того чтобы увидеть ещё более близкую точку В, приходится употребить ещё большее мускульное усилие для сосредоточения обоих глаз в точке В.

Таким образом, на рецепторное поле сетчатки глаза в область фовеа подается резкое изображение рассматриваемого объекта, при необходимости, и его деталей. Информация о цвете, форме, положении, размере и пр. передается в зрительную кору головного мозга.

¹ По теории Г. Гельмгольца, при рассмотрении предметов на различных расстояниях оптические параметры хрусталика меняются цилиарной мышцей или, как утверждают некоторые офтальмологи, перемещением внутриглазной жидкости, что для нас никакого значения не имеет, т.к. и то и другое приводит к изменению фокусного расстояния.

² Американский офтальмолог У. Бейтс на стыке XIX – XX веков сделал открытие, что изображение в человеческом глазе строится за счет изменения длины самого глаза.

Когда статья была уже готова, мне на глаза попала информация в «Большом психологическом словаре», которая подтверждает факт установки изображения в область фовеа.

«...Глаз – один из самых подвижных органов тела, не знающий покоя даже во сне (одна из фаз сна получила название «быстрые движения глаз»), но все главные функции движения глаз (ДГ) связаны, прежде всего, со зрительным восприятием, именно оно возлагает на ДГ определенные функции.

1. Первая очевидная функция ДГ состоит в том, чтобы перевести (установить) ретинальное изображение объекта, находящегося на периферии поля зрения, в центральную область сетчатки диаметром ок. 4° , называемую «центральной ямкой» (fovea centralis) или просто «фовеа» (лат. fovea – яма), которая обеспечивает высокую остроту зрения. В центре фовеа существует еще более оптимальный для восприятия участок – фовеола (уже не «ямка», а «ямочка» – размером ок. 1 в диаметре). Эта установочная функция отсутствует у тех видов животных, глаза которых не имеют центральной ямки (среди млекопитающих фовеа есть только у приматов, но она есть также у птиц, некоторых ящериц и даже рыб). Реализуется установочная функция с помощью т.н. баллистических, быстрых ДГ, которые принято называть «саккадическими ДГ» (скачкообразными). Когда мы осматриваем достаточно большой и сложно структурированный объект, изображение которого превышает размеры фовеолы (тем более фовеа, как, напр., страница книги), то приходится совершать много установочных движений. На длительных записях ДГ, совмещенных с осматриваемым объектом или сценой, можно наблюдать, что точки фиксации (остановок) концентрируются около наиболее информативных участков, что создает впечатление того, что глаза как бы ощупывают видимые объекты. Поэтому саккадические ДГ иногда называют поисковыми, обследующими, гностическими.

2. Если объект движется или же движется наблюдатель (или только его голова) относительно объекта, то возникает необходимость поддерживать ретинальное изображение примерно в одном положении, что и делают следящие и компенсационные ДГ (см. Нистагм). Можно сказать, что те и др. осуществляют функцию динамической фиксации, которая необходима даже тем видам животных, глаза которых не имеют фовеа. Поскольку при динамической фиксации глаза плавно подстраиваются под направление и скорость относительного движения объекта, то ДГ получили название следящих ДГ (син. плавные ДГ, медленные ДГ), чья минимальная скорость – ок. 5 угл. мин/с, что приблизительно соответствует пороговой скорости восприятия движения объекта; максимальная скорость – ок. 30–40 угл. град/с. Без специальной тренировки человек не способен произвольно вызывать медленные ДГ (вне ситуации слежения).

3. Особые задачи ставит перед ДГ бинокулярное зрение, для которого необходимо, чтобы ретинальное изображение объекта в правом и левом глазах попадало на корреспондирующие точки сетчатки. Из-за этого требования движения 2 глаз (как установочные, так и выполняющие функцию динамической фиксации) должны быть синхронными и содружественными, а зрительные оси (воображаемая линия, проходящая через центр зрачка, оптический центр глаза и центр фовеа; см. Линия взора) 2 глаз должны быть направлены в 1 точку. Если в порядке обеспечения указанных требований зрительные оси вращаются в одну сторону, то ДГ называются версионными движениями, если же зрительные оси сходятся или расходятся, т.е. двигаются в разные стороны, то ДГ относятся к типу вергентных движений. Вергентные ДГ требуются, когда новая точка фиксации находится либо дальше, либо ближе к наблюдателю» [1].

Существование фовеолы еще более усиливает эту функцию особенно для формирования единого представления при рассматривании больших объектов.

Однако есть и противники данной точки зрения.

«...Многочисленными исследованиями показано, что ДГ принимают активную роль в зрительном восприятии, участвуя в поиске и обнаружении объектов (стимулов), измерении и анализе пространственных свойств: форма, положение, размер, удаленность, скорость движения и др. Высказывались и возражения против этой точки зрения, опирающиеся на данные о том, что пространственные свойства могут оцениваться с достаточной точностью и без ДГ (напр., при очень краткой экспозиции, в условиях наблюдения последовательного образа или искусственно стабилизированного изображения на сетчатке). В свою очередь, защитники «глазодвигательной теории» приводят данные о викарных перцептивных действиях, а также обращают внимание на то, что ДГ особенно необходимы на ранних стадиях развития восприятия, а также в условиях наблюдения новых, малоизвестных объектов, когда еще не сформированы механизмы быстрого узнавания и анализа объектов» [1].

И все-таки, в зрительной коре мозга более 50% нейронов занято анализом информации, поступающей из области, соответствующей 10% поля зрения, принадлежащего его центральной части (центральной ямке), которая наиболее чувствительна к тонкой структуре и цвету изображения. В связи с этим зрительная система обладает возможностью переводить глаза с одной части поля зрения на другую, поскольку при любом фиксированном положении лишь небольшая часть поля зрения имеет высокое разрешение [2].

Действительно, у человека глаза все время находятся в движении, последовательно переходя с одного участка поля зрения на другой. Движение глаз складывается из скачков (саккад), которые обычно повторяются с частотой четыре – пять раз в секунду. Саккадическое движение происходит за счет активности наружных мышц глаз и, раз начавшись, продолжают до заданного положения без поправок во время движения (баллистическое движение). В целом характер движения глаз отражает систематический отбор внешней информации, основанный на осмысленной интерпретации поступающих данных [2].

При появлении в поле зрения стимула-цели глаза совершают скачок, в результате которого сетчатки изменяют свое положение так, что стимул-цель проецируется на каждой из них в область фовеа. Если после выполнения саккады стимул-цель все же не попадает в область фовеа, то следует корректирующая саккада. При возникновении акустического сигнала в одной из точек внешнего пространства глаза совершают саккаду и занимают такое положение, при котором направление на звук совпадает с направлением взора. При этом потенциально возможный зрительный образ источника звука будет проецироваться в область фовеа. При прикосновении к коже, например, груди, глаза поворачиваются так, что линия взора направлена на место раздражения. При этом фовеа глаз совмещаются с точкой прикосновения. Произвольные саккады реализуются и при осмотре зрительной сцены. Выбор стимула, представленного на константном экране, который подлежит уточнению с использованием фовеа, означает возбуждение такого командного нейрона, который обеспечит совмещение выбранной цели с фовеа при учете исходного положения глаз [3].

Соколов Е.Н. и Шмелев В.А. в книге «Нейробионика» пишут: «Саккадические движения глаз принадлежат к широкому классу баллистических движений, которые не контролируются на участке выполнения движения. Обратная связь при выполнении этих движений вводится через внешнюю среду после завершения элемента движения. Сходство с саккадическими движениями глаз обнаруживают саккадические движения головы, а также целенаправленные движения рук» [4]. (Думаю, что такой вид обратной связи характерен и для других функций головного мозга человека. Например, внутреннее или мысленное проговаривание читаемого текста, или внутреннее проговаривание мечтаний и мысленных размышлений, составление планов, планирование действий и принятие решений (Ав.)).

Далее они пишут, что при появлении нового зрительного стимула возникающие саккадические движения глаз представляют собой скачкообразный перенос взора на цель.

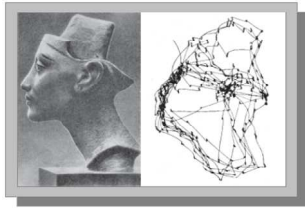


Рис. 5. Запись движений глаз при рассматривании фотографии скульптурного портрета Нефертити (Ярбус, 1965)

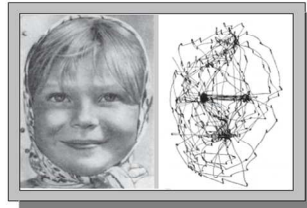


Рис. 6. Запись движений глаз при рассматривании фотографии девочки (Ярбус, 1965)

Если саккада не приводит к попаданию цели в область фовеа, генерируется корректирующая саккада, совмещающая цель с областью наилучшего видения в каждом из глаз. При перемещении цели в пространстве серия саккад выполняет функцию обратной связи, смещая проекцию цели в область фовеа, тем самым реализует слежение за целью [4].

Следует заметить, что управление саккадами глаз, движением головы, целенаправленными движениями рук, голосовых связок и пр. движениями осуществляется по сигналам цепочек командных нейронов, управляющих соответствующими перемещениями. При этом состояние цепочек командных нейронов запоминается (с помощью образования и укрепления новых связей между нейронами в нейросети) в процессе слежения за целью, при неоднократном повторении этих процессов в дальнейшем позволяет прогнозировать перемещение цели, движение рук или пр. изменения рецептивных полей в той или иной ситуации. Например, зная силу удара по мячу, мы свободно прогнозируем траекторию его полета и места приземления. Аналогично, общаясь с другом детства, членами семьи, коллегой, мы можем прогнозировать, о чем будет говорить тот или иной в той или иной обстановке. Прогнозирование – очень интересная тема, но вернемся к теме распознавания объектов.

Итак, при восприятии объекты сканируются движением глаз, и их изображение преобразуется в контурное.

На рис. 5 и 6 изображены фотографии головы Нефертити и девочки, а рядом записи движений глаз при свободном рассматривании фотографий. Оказывается, что взгляд движется по контуру объекта, при этом наибольшее количество остановок и повторных возвратов в тех местах контура, где больше всего тональных градаций, так как эти градации, вернее, их контраст и расположение, несут основную

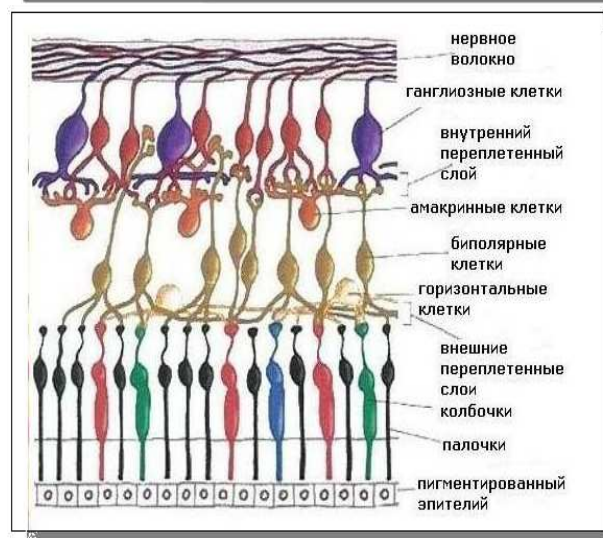
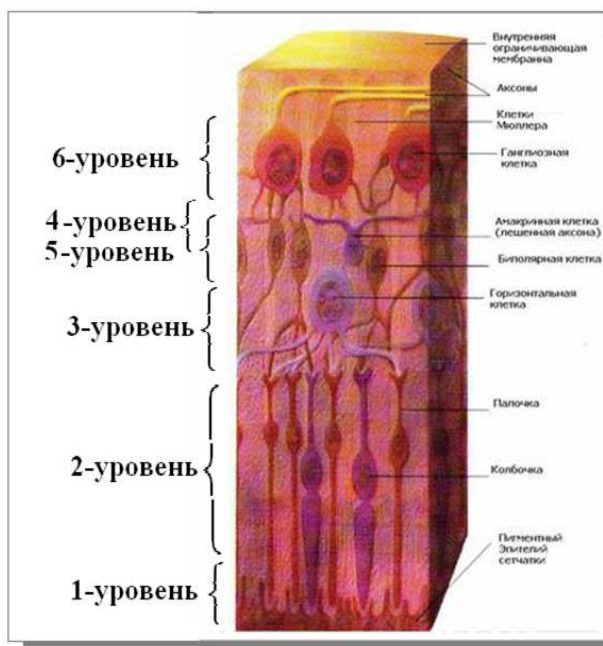


Рис. 7. Срез сетчатки глаза и его схематическое представление

информацию о форме и фактуре объекта.

Из вышеизложенного можно сделать вывод, что для распознавания лучше то изображение, которое имеет наибольшее количество тональных градаций, то, у которого наиболее четкий контур. Поэтому система преобразования тонального изображения в контурное имеет важное значение.

В биологических объектах первичная обработка информации осуществляется в сетчатке глаза на шести уровнях (рис.7). Первый уровень – пигментный эпителий сетчатки. Второй уровень – палочки и колбочки. Третий уровень – горизонтальные клетки. Четвертый уровень – амакринные клетки. Пятый уровень – биполярные клетки. Шестой уровень – ганглиозные клетки.

В целом назначение каждого уровня понятно. Первый и второй уровни – пигментный эпителий, палочки и колбочки – рецепторы восприятия видеoinформации. Палочки и колбочки не находятся в прямом контакте с корой головного мозга. Они контактируют с биполярными клетками, которые потом отправляют свои сообщения ганглиозным клеткам, аксоны которых составляют оптический нерв.

Третий уровень – горизонтальные клетки передают сообщения туда и обратно между клетками фоторецепторов, биполярным клеткам и друг другу. Горизонтальные клетки путем латерального торможения в окружающих областях останавливают диффузное распространение сигнала по сетчатке, которое могло бы возникнуть в связи с наличием широкого ветвления дендритов и аксонов в слоях сетчатки. Это важно для четкого выделения контрастных границ в зрительном образе. Четвертый уровень – амакринные клетки взаимосвязаны с биполярными клетками, ганглиозными клетками, а также друг с другом. Амакринные клетки располагаются во внутреннем зернистом слое сетчатки глаза, в плоскости внутреннего синаптического слоя сетчатки. Они образуют чрезвычайно сложную, соединенную щелевыми контактами сеть, а отросток, не покидая пределов сетчатки, ветвится в области синаптических контактов биполярных ганглиозных нейронов, они изменяют характер распространения возбуждения по нервным клеткам ганглиозного слоя. Число этих клеток особенно велико у животных с высокой остротой зрения, например, у птиц.

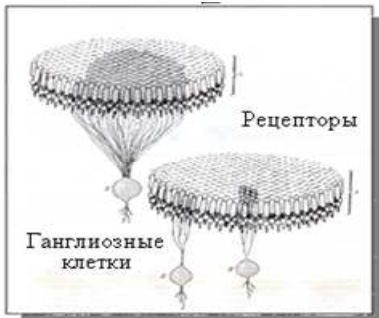


Рис. 8. Рецепторы сетчатки глаза

Третий и четвертый уровни играют важную роль в обработке визуальной информации на уровне сетчатки перед тем, как она передается в мозг для конечной интерпретации и служит для повышения соотношения чувствительность – детализация. Чем больше рецепторов присоединено к одной ганглиозной клетке, тем чувствительнее система, т.е. тем более слабые световые сигналы она сможет обнаружить (рис. 8). Сигнал на выходе ганглиозной клетки (острота зрения) будет тем сильнее, чем больше рецепторов к ней присоединено. При этом понятно, что повышение чувствительности приводит к потере детализации. Таким образом, третий и четвертый уровни обеспечивают чувствительность, четкость и контрастность воспринимаемого изображения.

Пятый уровень – биполярные клетки являются клетками промежуточного слоя сетчатки, передающими нервные импульсы от рецепторов ганглиозным клеткам. Шестой уровень – ганглиозные клетки очень разнообразны по характеру ветвления дендритов. Они расположены в сетчатке очень плотно, и их дендритные поля (рецептивные поля) перекрываются. На одну ганглионарную клетку может конвергировать (замыкаться) от одного до сотни биполярных нейронов. Через биполярные нейроны с одной ганглионарной клеткой может быть связано от единиц до десятков тысяч фоторецепторов (палочек и колбочек). В свою очередь, один фоторецептор через биполярные нейроны может быть связан с десятками ганглиозных клеток. Ганглиозные клетки завершают «трехнейронную рецеп-

торно-проводящую систему сетчатки»: фоторецептор – биполярный нейрон – ганглиозная клетка [5].

3. Аппаратно-нейронная модель сенсорного органа зрительной системы человека

Часть функций зрительной системы (первый, второй уровни) берут на себя видеокамеры, которые, как правило, используются в системах распознавания объектов. Третий – шестой уровни моделируются нейронной сетью. Полученное с видеокамеры тональное изображение объекта преобразуется в контурное, что позволяет значительно сократить объем необходимой для классификации информации.

Задача выделения контуров на изображении уже давно решается классическими алгоритмами. Среди них пространственное дифференцирование, функциональная аппроксимация, высокочастотная фильтрация. Общим для всех этих методов есть стремление рассматривать границу как область резкого перепада функции яркости изображения. Основным недостатком этих алгоритмов – относительно низкая скорость работы, которой будет недостаточно для обработки изображений большого размера при высокой частоте их поступления в реальном времени.

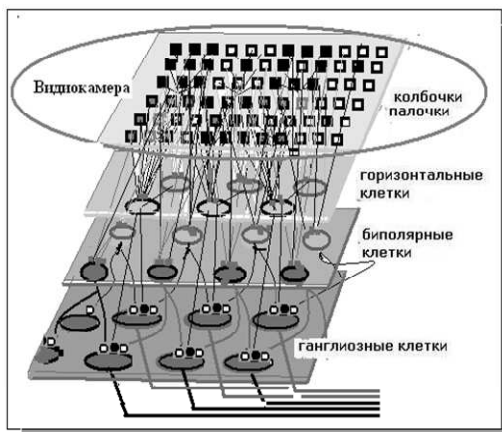


Рис. 9. Модель сенсорного органа зрительной системы человека (объемное представление)

Поэтому для выделения контуров изображения применен бионический подход, в котором обработка информации производится посредством нейронной сети. Такая сеть состоит из трех видов клеток, которые являются аналогом горизонтальных, биполярных и ганглиозных клеток (рис. 9). При физической реализации информация в сети обрабатывается на всех уровнях параллельно. В результате ожидается повышение быстродействия системы на несколько порядков.

На рис. 10 показана упрощенная схема модели сенсорного органа зрительной системы человека. Сигналы от соответствующего рецептора или групп рецепторов через горизонтальные клетки (на схеме не показаны) поступают на дендриты биполярной и ганглиозных клеток, усиливаясь положительным

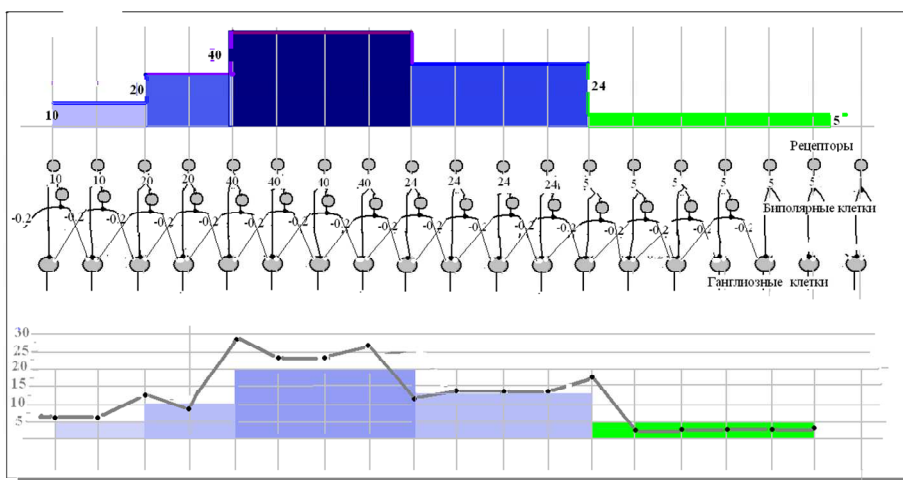


Рис. 10. Модель сенсорного органа зрительной системы человека (упрощенное представление)

весовым коэффициентом центрального дендрита, а сигналы, соответствующие соседним точкам изображения, тормозятся отрицательными весовыми коэффициентами боковых дендритов. Поступая в ганглиозную клетку, умноженные на соответствующие весовые коэффициенты сигналы суммируются и подаются на

выход. На выходе нейронной сети получается последовательность чисел, которые отвечают по определенному закону значениям кодов цветов входных точек изображения. Полу-

ченные на выходе данные изображены в виде графика (рис. 10). На графике видны резкие перепады полученной функции, совпадающие с резкими сменами цвета. Зафиксировав эти перепады, сравниваем их размер с заданным порогом. Превышение размера порога свидетельствует о наличии точки, принадлежащей контуру изображения.

3.1. Реализация модели сенсорного органа зрительной системы человека

Для проверки работы модели была создана система выделения контуров изображений. Благодаря полному параллелизму обработки каждой точки изображения, при аппаратной реализации этого подхода будет достигаться достаточно высокое быстродействие обработки в реальном времени входных теле- и видеоизображений. И так, в бионическом подходе считается, что каждая точка или условная группа точек воспринимаемого глазом изображения отвечает одному нейрону (ганглиозной клетке). Каждый такой нейрон имеет несколько дендритов, которые через биполярные клетки связаны отрицательными связями с соседними нейронами. Уровень возбуждения нейронов фактически соответствует качеству контура. В системе реализована гибкая система настройки параметров выделения контура:

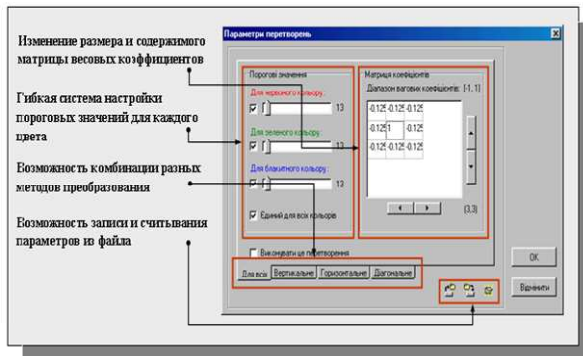


Рис. 11. Система настройки параметров выделения контуров

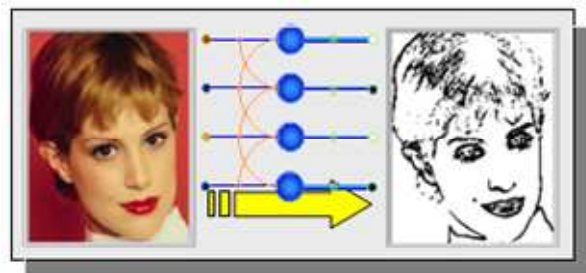


Рис. 12. Выделение контуров изображения

изменение размера матрицы весовых коэффициентов; гибкая настройка пороговых значений; комбинирование разных методов сканирования изображения; сохранение параметров настройки (рис. 11). Полученный контур, при оптимальной настройке системы, показан на рис. 12.

Таким образом, светочувствительные рецепторы преобразуют световой поток в нервные возбуждения (сигналы). Вот тут – то и возникают основные проблемы понимания функционирования зрительной системы, каким образом из сигналов, посылаемых по зрительным нервам, в неокортексе формируются образы, воспринимаемые зрением.

Наиболее вероятной моделью распознавания образов в высших слоях мозга человека является модель сравнения с эталоном. Это наиболее простой из всех способов распознавания образов. Распознавание осуществляется сравнением внешнего изображения с набором внутренних эталонов.

Рабочая гипотеза 2. Гипотеза о механизме распознавания образов в высших слоях неокортекса мозга человека. Механизмом распознавания образов в высших слоях неокортекса мозга человека является модель сравнения с эталоном. Данная модель обладает тем достоинством, что внешний образ сравнивается со всеми возможными эталонами одновременно. В процессе сравнения одновременно активируется некоторая часть эталонов, и тот, который реагирует на данный образ наиболее активно, и есть искомый объект.

В то же время модель имеет существенный недостаток, связанный с тем, что при изменении освещения, ориентации или размеров внешнего образа относительно эталона распознавания не произойдет. Эталон – внутреннее представление образа распознаваемого объекта, запомненного ранее в различных положениях. Следовательно, если распознавае-

мое изображение объекта приводить к стандартному освещению, размеру и точно совмещать с эталоном, то модель работает безукоризненно, учитывая, что при зрительном восприятии и концентрации внимания на объекте видеoinформация проецируется в центральную ямку сетчатки. При этом изображение приводится к стандартной освещенности, стандартному размеру центральной ямки глаза. Кроме того, осуществляется сканирование изображения с целью выявления наиболее информативных областей и, возможно, фиксации и сравнения расстояний между ними, то можно уверенно предположить, что зрительная система использует модель сравнения с эталоном.

Итак, мы считаем, что метод сравнения с эталоном активно применяется зрительной системой человека. В пользу этого предположения также свидетельствует и тот факт, что при рассмотрении (распознавании) изображения, например, очень сильно искаженной или перевернутой буквы или другого образа, встречающегося впервые, мы долго, с разных сторон, под разными углами зрения крутим его в руках. Но если мы узнали этот образ, то всегда в дальнейшем узнаем его легко, т.е. можно предположить, что искаженный образ сформировал новый эталон или был найден близкий эталон и зафиксированы связи, отвечающие за распознавание искаженного образа. Конечно, процесс обработки информации в слоях неокортекса не сводится только к сравнению объекта с эталоном. Этот процесс значительно сложнее. В нем на различных уровнях биологической нейронной сети, а в нашем случае – нейроподобной рецепторно-эффекторной растущей сети, осуществляются обработка, анализ, синтез и сравнение информации.

Модель сенсорного органа зрительной системы человека совместно с методологией обработки информации в рецепторно-эффекторных нейроподобных растущих сетях [6–8] положена в основу разработки виртуального робота «VITROM».

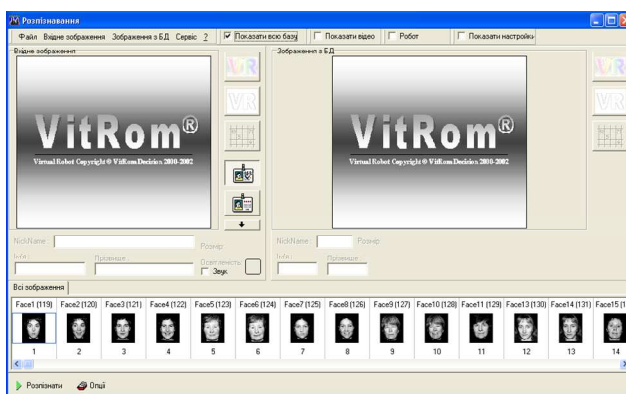


Рис. 13. Интерфейс виртуального робота «VITROM»

4. Виртуальный робот «VITROM»

Интерфейс виртуального робота «VITROM» показан на рис. 13. При отсутствии объекта распознавания в зоне видимости видеокamеры («глаза робота») робот призывает объект подойти к нему. При появлении объекта в зоне видимости робот пытается распознать объект. Если объект находится в зоне видимости, но на значительном удалении робот приглашает подойти ближе и пытается распознать объект. Если объект известен роботу, то он здоровается и называет объект по имени и отчеству.

Если робот не «знает» объект или ошибся, то он предлагает объекту представиться, запоминает информацию и его изображение.

Для проверки правильности распознавания изображений лица человека виртуальный робот «VITROM» ознакомлен с базой «Yale FaceIMAGES_Data» и другими 600-ми изображениями. Распознавание практически безошибочное.

5. Выводы

В результате изучения существующей нейрофизиологической литературы сформулированы гипотезы «О приведении распознаваемых изображений к одному размеру в области фовеа» и «О механизме распознавания образов в высших слоях неокортекса мозга человека». На базе рабочих гипотез созданы и реализованы модель и нейронная сеть сенсорного органа зрительной системы человека, подтвердившие правомерность гипотез. Модель сен-

сорного органа зрительной системы человека совместно с методологией обработки информации в рецепторно-эффекторных нейроподобных растущих сетях положена в основу создания программной системы распознавания лиц, которая показала высокую эффективность распознавания. При аппаратной реализации системы информация в рецепторно-эффекторных нейроподобных растущих сетях обрабатывается на всех уровнях параллельно. В результате ожидается повышение быстродействия системы на несколько порядков.

Программная реализация системы демонстрировалась на выставке в Пекине 2000г., в Ганновере на выставках СеВІТ 2000–2002 гг., в павильонах ВДНХ Украины 2000–2004 гг.

СПИСОК ЛИТЕРАТУРЫ

1. Движения глаз. Большой психологический словарь [Электронный ресурс]. – Режим доступа: wiki.myword.ru/indtx.php.
2. Линдсей П. Переработка информации у человека (Введение в психологию) / П. Линдсей, Д. Норман; под ред. А.Р. Лурия. – М., 1974. – С. 549.
3. Соколов Е.Н. Нейронные механизмы саккадических движений глаз // Вопросы психологии. — 1980. – № 3. – С. 64 – 77.
4. Соколов Е.Н. Нейробионика. Организация нейроподобных элементов и систем / Е.Н. Соколов, В.А. Шмелев. – М.: Наука, 1983. – 279 с.
5. <http://ru.wikipedia.org/wiki>.
6. Ященко В.А. Рецепторно-эффекторные нейроподобные растущие сети – эффективное средство моделирования интеллекта. I / В.А. Ященко // Кибернетика и системный анализ. – 1995. – № 4. – С. 54 – 62.
7. Ященко В.А. Рецепторно-эффекторные нейроподобные растущие сети – эффективное средство моделирования интеллекта. II / В.А. Ященко // Кибернетика и системный анализ. – 1995. – № 5. – С. 94 – 102.
8. Yashchenko V.A. Receptor-effector neural-like growing network – an efficient tool for building intelligence systems / V.A. Yashchenko // Proc. of the second international conference on information fusion, (July 6–8, 1999, Sunnyvale Hilton Inn, Sunnyvale, California, USA). – California, USA, 1999. – Vol. II. – P. 1113 – 1118.

Стаття надійшла до редакції 26.12.2011