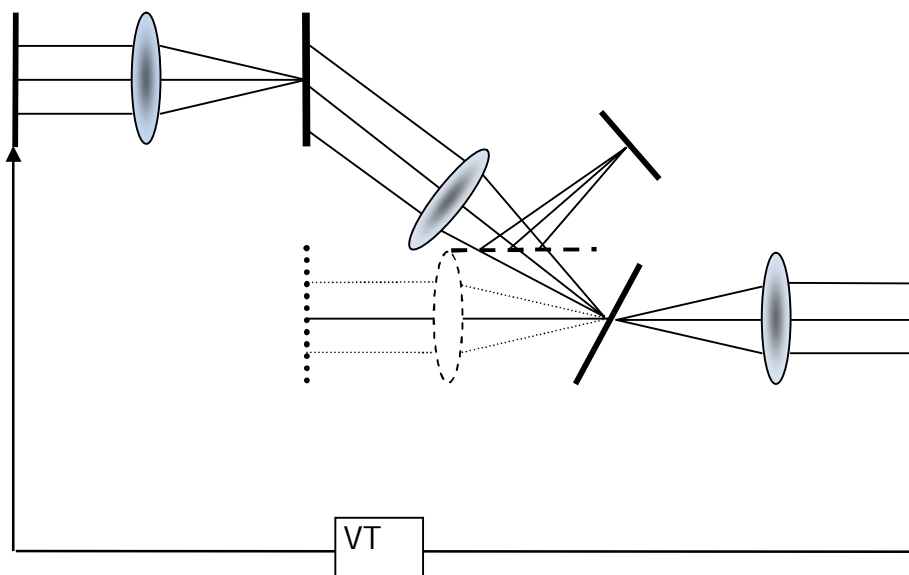


В.Н.Васильев, А.В. Павлов
ОПТИЧЕСКИЕ ТЕХНОЛОГИИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Том 1.

Основы оптических информационных технологий и искусственных
нейронных сетей



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

В.Н. Васильев, А.В.Павлов
ОПТИЧЕСКИЕ ТЕХНОЛОГИИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
Учебное пособие

Том 1.

Основы оптических информационных технологий
и искусственных нейронных сетей

 **УНИВЕРСИТЕТ ИТМО**

Санкт-Петербург

2017

В.Н.Васильев, А.В.Павлов, Оптические технологии искусственного интеллекта / Учебное пособие. Изд.4-е, дополненное. В 2-х т. Т.1. – СПб: Университет ИТМО, 2017. – 80 с.

В пособии представлены методические материалы по курсу «Оптические технологии искусственного интеллекта». Кратко изложены базовые сведения из области искусственного интеллекта, теории искусственных нейронных сетей, теории нечетких множеств, парадигмы когнитивной системы, включая начала теории динамического хаоса. Рассмотрены вопросы применения оптических технологий для реализации ряда моделей искусственных нейронных сетей, нечетких логик и когнитивных систем.

Для бакалавров, обучающихся по направлению подготовки 12.03.03 «Фотоника и оптоинформатика» по профилям подготовки: "Оптические и квантовые технологии передачи, записи и обработки информации", "Компьютерная фотоника".

Рекомендовано к печати Ученым советом факультета фотоники и оптоинформатики, протокол № 2 от 1 марта 2017 г.



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО – становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО, 2017
©В.Н.Васильев, А.В.Павлов, 2017

«Феномен информации есть многостадийный, необратимый процесс становления структуры в открытой неравновесной системе, начинающийся со случайного запомненного выбора, который эта система делает, переходя от хаоса к порядку, и завершающийся целенаправленным действием согласно алгоритму или программе, отвечающим семантике выбора».

И.В.Мелик-Гайказян, «Информационные процессы и реальность», М., 1997.

ОГЛАВЛЕНИЕ

Том 1	Стр.
Список обозначений и сокращений	4
Лекция 1. Вводная	5
Предмет нашего изучения и подходы к проблеме	5
Выбор подхода к построению системы ИИ и парадигмы ИИ	9
Лекция 2. Базовые математические операции	17
Преобразование Фурье	18
Интегральные операции, реализуемые через преобразование Фурье	24
Свойства операций свертки и корреляции	25
Теорема Котельникова	27
К вопросу биологической обоснованности использования преобразования Фурье в искусственных нейронных сетях	30
Лекция 3. Распознавание образов. Оптические методы распознавания	32
Лекция 4. Основы парадигмы искусственных нейронных сетей	42
Лекция 5. Сети однонаправленного распространения	52
Лекция 6. Тема 6.1. Искусственные нейронные сети как динамические системы.	62
Двунаправленная ассоциативная память	62
Сети Хопфилда – сети полной связности	64
Лекция 6. Тема 6.2. Самообучение в нейронных сетях. Сеть Кохонена	68
Лекция 8. Концепция ART	72
Кафедра фотоники и оптоинформатики	80

Список принятых сокращений и обозначений.

ВИ	Вычислительный интеллект
ГК	Голографический коррелятор
ГРС	Голографическая регистрирующая среда
ДЭ	Дифракционная эффективность
ИИ	Искусственный интеллект
КП	Когнитивная психология
КмП	Компьютерная парадигма в ИИ
КС	Когнитивная система
ЛМ	Лингвистическая модель
ЛП	Лингвистическая переменная
МЛ	Многозначная логика
НЛ	Нечеткая логика
НзЛ	Нечеткозначная логика
НмЛ	Немонотонная логика
НМ	Нечеткое множество
НС	Нейронная сеть
ОНС	Оптическая нейронная сеть
СП	Синаптическая пластичность
ТНМ	Теория нечетких множеств
AI	Artificial Intelligence
ART	Adaptive Resonance Theory (теория адаптивного резонанса)
FST	Fuzzy Sets Theory (теория нечетких множеств)
LV	Linguistic Variable (лингвистическая переменная)
NN	Neural Network (нейронная сеть)
WTA	Winner Takes All (победитель забирает все)

Лекция 1. Вводная

Предмет нашего изучения и подходы к проблеме

Как следует из самого названия предмета, искусственный интеллект (ИИ) относится к категории наук об искусственном, т.е. созданном (или создаваемом) человеком. Этот тип наук вторичен в том смысле, что не только имеет некий естественный прототип (для ИИ это биологический интеллект (БИ), но и само развитие таких наук в значительной степени зависит от успехов наук, изучающих эти естественные прототипы. В частности, при возникновении такой типичной для истории развития любой науки кризисной ситуации как исчерпанность парадигмы*, новая парадигма вырабатывается не внутри самой науки об искусственном, как это происходит в естественных науках (например, кризис классической парадигмы в физике рубежа XIX–XX веков), но индуцируется извне осмыслением результатов, полученных науками о естественном прототипе.

Биологический интеллект изучается в рамках ряда наук, каждая из которых уделяет преимущественное внимание одной из сторон этого многогранного явления: анатомии, нейрофизиологии, психологии и др.

Анатомия и нейрофизиология уделяют основное внимание изучению мозга как материального носителя интеллекта. Успехи, достигнутые в этих науках, индуцировали, в частности, возникновение и развитие такой мощной парадигмы в ИИ как искусственные нейронные сети (ИНС).

Изучение информационных процессов сосредоточено в значительной степени в рамках направлений, объединенных термином «когнитивные науки». Для нас интерес представляют следующие направления:

Бихевиоризм – направление в психологии, которое отрицало сознание как предмет научного исследования и сводило психику к различным формам поведения, понимаемого как совокупность реакций организма на стимулы внешней среды. Как реакция на характерное для бихевиоризма отрицание роли внутренней организации психических процессов в конце 50-х – начале 60-х годов XX в. возникло такое направление, как *когнитивная психология (КП)*.

Гештальт психология. Немецкое слово «гештальт» означает целостную систему, акцентируя внимание именно на целостности системы. Это направление акцентирует внимание на включенности носителя интеллекта в окружающий его мир, с которым он составляет единую систему. Отсюда следует признание пассивной роли субъекта в познании мира – интеллект пассивно отражает мир, но не выдвигает гипотез и не ставит над окружающим миром опыты. Иногда в русском языке это научное направление определяют также термином «экологическая психология».

* Парадигма – (др.-греч. Παράδειγμα – «пример, модель, образец») – совокупность признаваемых на данном этапе истинными постулатов и предпосылок, определяющих рамки исследований и критерии научности.

Предмет когнитивной психологии (КП): восприятие, распознавание образов, внимание, память, воображение, речь, психология развития, мышление и решение задач, человеческий и искусственный интеллект.

Основной метод КП: анализ микроструктуры того или иного психологического процесса.

Первоначально главной задачей КП виделось изучение преобразований сенсорной информации от момента попадания стимула на рецепторные поверхности до получения ответа. При этом исследователи исходили из аналогии между процессами переработки информации у человека и в вычислительном устройстве. Такой подход известен под названием «компьютерной метафоры» в КП. Были выделены многочисленные структурные составляющие (блоки) познавательных и исполнительных процессов, в том числе кратковременная память и долговременная память. Эта линия исследований, столкнувшись с серьезными трудностями в связи с увеличением числа структурных моделей частных психических процессов, привела к выработке нового понимания когнитивной психологии как направления, задачей которого является доказательство решающей роли знания в поведении субъекта. Такое понимание предмета КП поставило в центр этой науки вопрос об организации знания в памяти субъекта, в том числе о соотношении вербальных и образных компонентов в процессах запоминания и мышления. И, соответственно, определило роль КП как одного из важнейших естественно-научных оснований ИИ. Поэтому в настоящее время термин КП используется в широком смысле.

Для КП характерно понимание интеллекта (и его носителя) как активно действующего агента, познающего окружающий мир посредством активного в него вторжения, строящего субъективную модель мира, выдвигающего гипотезы об устройстве мира и проверяющего их на опыте.

Существует множество определений интеллекта [1-3], отражающих ту или иную сторону этого многогранного явления, например:

- интеллект как способность решения задач на основе манипуляций над символами – на основе этого определения зародилось символическое (логическое) направление в искусственном интеллекте;
- интеллект – прежде всего способность к обучению – на основе этого определения сформировалась *коннекционистская парадигма*, развившаяся в парадигму *искусственных нейронных сетей* (ИНС);
- интеллект – способность оперативной адаптации к динамической, быстроменяющейся, сложной среде – на основе этого определения сформирована *эволюционистская парадигма* в ИИ, включающая такие направления как эволюционные вычисления и генетические алгоритмы.

Из других определений, даваемых по отдельному признаку, вышли такие классические направления в комплексной проблеме ИИ, как машинное восприятие, анализ сцен, машинный перевод и т.п.

Реальный феномен интеллекта может быть определен через совокупность его способностей, а именно (цитируется по В.К.Финну [4]):

1. *Способность выделять существенное в наличных знаниях, т.е. упорядочивать их;*
2. *Способность к целеполаганию и планированию поведения: порождение последовательности «цель → план → действие»;*
3. *Способность к отбору знаний, релевантных задаче;*
4. *Способность искать следствия из имеющихся знаний, т.е. способность к рассуждению, которое может содержать как правдоподобные выводы, используемые для выдвижения гипотез, так и достоверные выводы;*
5. *Способность к аргументированному принятию решений, использующему упорядоченные знания и результаты рассуждений;*
6. *Способность к рефлексии: оценки знаний и действий;*
7. *Наличие познавательного любопытства: познающий субъект должен быть способен порождать вопрос «что такое?» и искать на него ответ;*
8. *Способность и потребность находить объяснение (не обязательно дедуктивное) как ответ на вопрос «почему?» (реализация абдуктивного инстинкта в смысле Ч.С. Пирса);*
9. *Способность к синтезу познавательных процедур, образующих эвристику решения задач и рассмотрения проблем, например, реализация взаимодействия индукции, аналогии и абдукции (с учетом фальсификации выдвигаемых гипотез посредством поиска контрпримеров) с последующим применением дедукции;*
10. *Способность к обучению и использованию памяти;*
11. *Способность к рационализации идей: стремление уточнить их как понятия;*
12. *Способность к созданию целостной картины относительно предмета мышления (формирование, по крайней мере, приближенной «теории» предметной области);*
13. *Способность к адаптации в условиях изменения жизненных ситуаций и знаний (что означает коррекцию «теорий» и поведения).*

Следует иметь в виду, что приведенная характеристика описывает лишь идеальный тип и выражает существенные черты феномена интеллекта, отнюдь не всегда, увы, присущие конкретному индивиду.

Под термином искусственный интеллект в настоящее время понимается как научное направление, так и конечный продукт, на получение которого направлены усилия в рамках этого направления.

Г.С.Осипов, президент Российской ассоциации искусственного интеллекта, определяет ИИ как научное направление следующим образом:

«Наука под названием «искусственный интеллект» входит в комплекс компьютерных наук, а создаваемые на ее основе технологии относятся к информационным технологиям.

Задачей этой науки является обеспечение разумных рассуждений и действий с помощью вычислительных систем и иных искусственных устройств. На этом пути возникают следующие главные трудности:

- а) в большинстве случаев до получения результата не известен алгоритм решения задачи. Например, точно неизвестно, как происходит понимание текста, поиск доказательства теоремы, построение плана действий, узнавание изображения.

б) искусственные устройства не обладают достаточным уровнем начальной компетентности. Специалист же добивается результата, используя свою компетентность (знания и опыт).

Это означает, что ИИ представляет собой экспериментальную науку. Экспериментальность ИИ состоит в том, что создавая те или иные компьютерные (или иные) представления и модели, исследователь сравнивает их поведение между собой и с примерами решения тех же задач специалистом, модифицирует их на основе этого сравнения, пытаясь добиться лучшего соответствия результатов.»

Далее мы рассмотрим основные парадигмы ИИ и увидим, что приведенное определение в значительной степени соответствует компьютерной парадигме в ИИ, поскольку не просто относит ИИ к компьютерным наукам, но и предполагает её эвристический характер.

Продукт ИИ: искусственные системы, осуществляющие конструктивное приближение вышеописанных способностей. Существует понятие «сильного ИИ» и «слабого ИИ».

Под первым понимается искусственная система, не отличимая по своим интеллектуальным способностям от человека. В качестве критерия и конструктивного определения «сильного ИИ» известен мысленный эксперимент, известный как тест Тьюринга. Можно рассматривать эту задачу как задачу приближения сингулярности в её сверхинтеллектуальном понимании.

Под вторым – система решающая те или иные задачи из ограниченного круга, например, распознавание образов, понимание речи и т.п.

Поскольку предметом настоящего курса являются технологии, то есть конкретные технические методы решения конкретных практических задач, то в рамках настоящего курса используем упрощенное определение и *под системой ИИ будем понимать автономную систему, решающую задачу целеполагания, самостоятельной ориентации и принятия решений, ведущих к достижению цели, в сложной, непредсказуемой обстановке во взаимодействии с другими системами ИИ. При этом как сама цель, так и критерии ее достижения могут быть определены нечетко, неоднозначно, допускать многовариантность интерпретации и смену приоритетов в процессе жизнедеятельности.*

Как следует из принятого определения ИИ, основным критерием "интеллектуальности" технических систем можно принять их способность к выработке решений на основе воспринимаемой информации. Соответственно, эти две взаимосвязанные задачи - восприятие информации из окружающего мира и выработка решения и составляют научную и техническую суть проблемы создания ИИ. Принятое определение ИИ также показывает, что речь не идет о попытке создать техническими средствами полный аналог интеллекта биологического, но лишь о заимствовании от биологических прототипов отдельных идей,

принципов организации и функционирования при создании технических систем, решающих определенный круг задач.

Возможно ли создание ИИ, полностью идентичного интеллекту биологическому? Поскольку понятие ИИ часто ассоциируется со столь излюбленной фантастами темой бунта ИИ против своего создателя, то напомним, что для создания ИИ, полностью идентичного биологическому, прежде необходимо познать *что* есть интеллект биологический, т.е. человек должен познать самого себя. Это уже вопрос не технический, а философский и теологический и те философские системы, которые допускают познание человеком самого себя, признают это вершиной познания, завершением человеческой истории в ее нынешнем виде.

Возвращаясь к технической стороне вопроса (в надежде, что история еще не завершается), в очерченных рамках проблемы ИИ необходимо, помимо прочих, первоочередное решение двух взаимосвязанных задач:

1. *Выбор теоретического подхода к построению системы ИИ;*
2. *Выбор физических механизмов и технических средств их реализации, адекватных выбранному подходу.*

Возможные методы решения этих задач и рассмотрены в данном курсе.

Выбор подхода к построению системы ИИ и парадигмы ИИ

ИИ как направление зародился во второй половине 50-х годов в ходе общекибернетических дискуссий о природе информационных процессов мозга. Бурный прогресс 50-70х годов в развитии кибернетики и ЭВМ сопровождался своеобразной эйфорией, когда казалось, что любая задача может быть представлена в форме машинной программы и сосчитана. Проблема представлялась лишь в том, чтобы разработать достаточно подробный и адекватный алгоритм, увеличить объем памяти и быстродействие ЭВМ. Этот подход, основанный на логическом выводе в символической форме и точных вычислениях, известен в истории ИИ под названием *компьютерной парадигмы (КМП)* [1,2].

Основные черты *КМП* – символьное представление информации и ее обработка алгоритмируемыми логико-комбинаторными методами. Символ здесь понимается только как знак алфавита, не имеющий ни семантики, ни прагматики.

Именно в это время на страницы печати выплеснулась дискуссия на тему «может ли машина думать?», а понятие «тест Тьюринга» стало знакомо каждому школьнику.

Однако, по мере возрастания сложности алгоритмов, увеличения объема памяти и вычислительной мощности ЭВМ стало ясно, что есть задачи, требующие для своего решения принципиально иного подхода [3]. Многие виды интеллектуальной деятельности в рамках *КМП* либо не моделируются вовсе, либо моделируются с недостаточной эффективностью, которая в рамках алгоритмического подхода неустраима в принципе. На повестку дня встал вопрос: «Верна ли

исходная посылка *КМТ* о возможности рассмотрения безотносительно к их физическому носителю? Иными словами – возможно ли моделирование информационных процессов мозга в терминах формальных систем?

Ответ на этот вопрос дали результаты психологических и нейрофизиологических исследований работы мозга, которые показали, что ряд задач безнадежно медленный по сравнению с ЭВМ мозг решает быстрее и эффективнее прежде всего потому, что решает принципиально по другому, чем ЭВМ [5-8]. Круг задач, решаемых интеллектом, весьма условно (в меру нашего понимания) может быть разбит на два класса:

- *задачи, требующие строго-логических методов решения и*
- *трудноформализуемые, комбинаторно-комплексные, нечетко определенные, не алгоритмизируемые и не вербализируемые.*

Такое разграничение в определенной степени соответствует известному функциональному разделению полушарий мозга - левое полушарие по преимуществу решает задач из первого класса, правое - второго. Подчеркнем еще раз, что это разделение весьма условно и соответствует скорее уровню нашего понимания, чем реальности, так как даже лево-полушарные информационные процессы у человека устроены совсем не так, как у ЭВМ – логико-алгебраические процессы лишь их часть [1]. В настоящем курсе сосредоточимся на проблемах применения ОИТ в рамках второго класса задач, т.е. задач, требующих для своего решения иной, нежели компьютерная, парадигмы.

Принципы информатики мозга, не учитываемые в рамках *КМТ*

1. Мозг здорового бодрствующего человека – предельно неустойчивая хаотическая система.
2. Функциональная асимметрия полушарий – различна нейрофизиологическая природа протекающих процессов. Правое полушарие – процессы не алгоритмизируемые, не вербализируемые. Но и левополушарные информационные процессы у человека протекают не так, как у ЭВМ – логико-алгоритмические процессы – лишь часть левополушарных процессов. Для человека стремление к целостности и устойчивости информационных структур первично по отношению к логике.
3. Характерная особенность ряда информационных процессов мозга – их невоспроизводимость и невербализуемость, т.е. невозможность их объективации, формального описания, с помощью которого было бы можно их повторить и передать другому.
4. Несимвольное, аналоговое представление информации. На концептуальном уровне преобразования информации непрерывные, но дискретизируемые в целях их объективизации и передачи.
5. Образность мышления. Внутренние репрезентации информации суть картины нейронной активности коры головного мозга. ***NB!*** Многие исследователи обращают внимание на наличие сильнейшего

внутреннего противоречия, скрытого в часто употребляемом словосочетании «моделирование образов».

6. Работа с отношением сходства (толерантности), а не с отношением тождества. Наличие эффектов размывания и обострения образов.
7. Распределенность информации – даже локальная информация хранится и считывается глобально.
8. Параметричность модели – информационные свойства модели должны зависеть не только от архитектуры, но и от параметров, модификация которых позволяет менять свойства модели без перенастройки связей.

Два типа мышления: научное и обыденное.

Научное мышление логично, для него характерно стремление к абстрагированию.

Обыденное мышление – для него важнее целостность и устойчивость внутренней картины мира, нежели ее логичность. Характерно стремление к заострению образов. Именно стремлением к устойчивости объясняются такие невозможные с точки зрения строгой логики, но тем не менее всем знакомые явления, как «православный коммунист» или люди, до дрожи ненавидящие конкретных чиновников, но дружно голосующие за тех же самых чиновников, объединившихся в партию власти.

Два типа мировосприятия: мужское и женское.

Мужской тип – познание в сущностях.

Женский тип – познание в отношениях (термин отношение используется здесь в математическом смысле).

Именно различием в типах мировосприятия объясняется то, что женщины с точки зрения мужчин постоянно болтают языком вместо того, чтобы делать дело, а мужчины с точки зрения женщин столь невнимательны и примитивны. В разговорах женщины строят отношения, т.е. познают мир, строят его внутреннюю картину. Напротив, мужчинам важно дело (т.е. сущность), а не разговоры вокруг дела (отношения).

Два типа процессов мышления: быстрые и медленные.

Быстрые – работа с уже существующими структурами.

Медленные – создание структур (подготовка новых идей, завершающаяся озарением).

Поэтому быстрота и живость ума отнюдь не свидетельствуют о выдающихся интеллектуальных способностях. Многие выдающиеся ученые и мыслители были тугодумами и испытывали из-за этого большие трудности во время учебы. Часто (но не всегда!) человек задумывается над казалось бы очевидными вещами потому, что стремится увидеть то, что сокрыто в глубине под лежащим на поверхности ответом. Многих это раздражает: «чего там думать, когда и так все ясно!». Но именно из размышлений над казалось бы всем очевидными вещами и рождаются великие открытия и прорывы мысли (Ньютон и яблоко).

Переход от работы с данными к работе со знаниями.

Проф. Д.А.Поспелов характеризует знания следующими свойствами:

1. Внутренняя интерпретируемость знаний (понятность знания его носителю).
2. Структурированность знаний. Информационные единицы должны обладать гибкой структурой. Принцип «матрешки» – рекурсивная вложимость знаний. Возможность произвольного установления и перенастройки отношений (включения) между информационными единицами.
3. Связность. Отношения между элементами: структурные, функциональные, каузальные и семантические. Структурные задают иерархию, функциональные задают процедурную информацию, позволяющие находить одни элементы через другие, каузальные задают причинно-следственные связи, семантические охватывают все остальные виды отношений.
4. Ассоциативность знаний – наличие семантической метрики в сфере знаний. Отношение релевантности на множестве информационных единиц характеризует ситуационную близость элементов (силу ассоциативной связи). Позволяет находить знания, близкие к уже найденным.
5. Активность знаний – наличие у знаний побуждающей и направляющей функции, что фактически превращает знания в квазипотребности. Актуализации тех или иных действий способствуют имеющиеся в системе знания.

Структура индивидуального знания может быть формально представлена в виде двух математических моделей: семантической ассоциативной сети и семантической пропозициональной сети.

Семантической ассоциативной сети соответствует интуитивный когнитивный стиль и полнезависимость, «текущий интеллект». Характерна метафоричность языка, связь компонентов сложного предложения по типу сочинения. Описывает знания новичка в данной области.

Семантической пропозициональной сети соответствует рациональный когнитивный стиль, полнезависимость, высокий коэффициент абстракции, аналитическая способность. Логические отношения, сукцессивная организация речи (текста). Язык неметафорический, указания на существенные детали, связь компонентов сложного предложения по типу подчинения. Описывает экспертные знания.

Актуализация той или иной сети определяется ситуацией.

Попытка ответа на поставленные выше вопросы с учетом результатов, полученных в рамках анатомии, нейрофизиологии и когнитивной психологии привела к формированию в 90-е годы XX века новой парадигмы ИИ – парадигмы вычислительного интеллекта (ВИ). ВИ объединяет ряд научных направлений (или микропарадигм), основные из которых следующие (3 источника и 3 составных части ВИ):

- *Коннекционизм* базируется на результатах нейрофизиологических исследований мозга и относится к категории *biologically inspired models*, наиболее мощное направление – ИНС. Может рассматриваться как теория «hardware» ИИ. Основная идея заключается в том, что если создать искусственную конструкцию, подобную естественной (мозгу), то и работать (мыслить) она должна подобным образом.
- *Теория Нечетких Множеств (ТНМ)* развивает традиции применения математической логики к описанию реальных процессов мышления и представляет попытку языком математики описать присущую человеческому мышлению и мировосприятию неполноту, нечеткость и противоречивость. Основатель ТНМ – уроженец г.Баку Лотфи Заде (Lotfi Zadeh). Легенда, поддерживаемая самим Заде, гласит, что ТНМ началась со спора – Заде взялся языком математики описать красоту своей жены. Это направление может быть отнесено к категории «Software» ИИ. Отметим, что понятие нечеткости, лежащее в основе ТНМ, – вариант более общего НЕ-фактора *неоднозначность* [19].
- *Эволюционные вычисления* и генетические алгоритмы. Попытка применить знания, полученные в генетике, к созданию искусственных саморазвивающимся систем, адаптирующихся в процессе жизнедеятельности к изменяющимся внешним условиям.

Мы рассмотрим оптические технологии только для двух первых составляющих парадигмы ВИ: ИНС и ТНМ. Объединение этих подходов дает синергетический эффект за счет сочетания вычислительной мощи и обучаемости ИНС с гибкостью ТНМ, известно как «*Мягкие вычисления*» («Soft Computing», термин введен Л. Заде). Основания к применению оптических технологий в этих рамках дают следующие аналогии между фундаментальными физическими явлениями и атрибутами ВИ:

ИНС

Атрибут парадигмы	Физическое явление
Полная связность	Явление дифракции Связи в свободном пространстве
Обучаемость вместо программирования	Явление интерференции и техника голографии
Ассоциативность мышления	Ассоциативность отклика голограммы

ТНМ

Атрибут парадигмы	Физическое явление
Фаззификация (увеличение нечеткости) обрабатываемой информации	Явление дифракции
Нечеткое отношение множеств	Явление интерференции и техника голографии
Отображение множества на нечетком отношении	Дифракция освещающего пучка на голограмме

Необходимо также иметь в виду, что интеллект невозможен без сенсоров (получение информации о внешнем мире) и исполнительных органов (активное вторжение в мир для постановки опытов). Поэтому Дж.Гиббсоном было введено понятие *аффорданса* как единства восприятия и действия. Признание неразрывности восприятия и мышления нашло свое отражение в предложенном Л.Заде «сдвиге парадигмы» – от парадигмы ВИ к парадигме вычислительного восприятия (ВВ).

Дальнейшее развитие ИИ привело к формированию в самые последние годы такого понятия как «*Когнитивная система*» (КС) [20]. Как следует из названия, в понятии КС акцент делается на познание – именно способность познавать есть главное отличие КС от традиционных информационных систем. КС – это система, которая способна познавать свое окружение и влиять на него. КС реализует переходы «данные → информация → знание». Особенности познавательных процессов КС суть следующие [20]:

- *познание базируется как на имеющемся знании, так и на результатах наблюдения;*
- *познавательный процесс порождает не выводы, а гипотезы (абдукция вместо дедукции);*
- *система в результате познания меняет свое знание, а следовательно, и свою способность переработки информации;*
- *познание – целенаправленное действие, предполагающее рефлексивное управление.*

Здесь необходимо упомянуть, что еще в 30-х годах XX века отечественным физиологом П.К.Анохиным была сформулирована и в дальнейшем развивалась его школой концепция «*Функциональной системы*» (ФС), во многом предвосхитившая парадигму КС. Многие вопросы в парадигме ФС еще в середине XX века были проработаны глубже и шире, чем в КС. В частности, именно П.К.Анохиным, еще за 13 лет до появления самого термина «кибернетика» было введено понятие обратной связи. Именно П.К.Анохин определил полезный результат, на достижение которого направлена деятельность системы, как системообразующий фактор. Именно П.К.Анохин обратил внимание и исследовал способность к предсказанию как важнейший атрибут интеллекта и механизмы ее реализации. Список приоритетов может быть продолжен. К сожалению, значительную часть своих сил и времени П.К.Анохин был вынужден тратить на борьбу с ревнителями чистоты марксистско-ленинского учения, благодаря которым Россия в очередной, увы, раз отдала свой научный приоритет западным исследователям. Рассмотрению ключевых моментов ФС посвящена отдельная лекция.

В процессе изучения этого курса, на конкретных примерах мы увидим, что именно техника голографии в наибольшей степени соответствует атрибутам этих парадигм. Наличие ряда глубоких аналогий между

свойствами человеческого интеллекта и оптической голографии было отмечено еще на первых этапах развития голографии основателями этого направления Д.Габором и Ю.Н.Денисюком. Признание актуальности этих аналогий специалистами по ИИ и КП нашло свое отражение в формировании «голографической парадигмы» (ТТ) в рамках этих научных направлений. Следует отметить, что предпосылки для формирования ТТ в КП были заложены еще в 1906г. немецким врачом А. Гольдшейдером, предложившим рассматривать механизмы восприятия и памяти как резонансные взаимодействия между волновыми фронтами, которые создаются при поступлении сенсорных потоков в кортикальные области мозга. Позже, познакомившись с голографией, американский психолог Карл Прибрам выдвинул гипотезу о голографическом устройстве мозга, которая объединяет логическое и психическое начала интеллекта [13]. Согласно голографической метафоре, актуально переживаемую психическую реальность можно рассматривать как стоячую волну, возникающую в результате функционирования организма как целостной системы, объединяющей сенсоры, процессоры и исполнительные органы.

Голографическая парадигма получила признание не только среди специалистов по ИИ, но и среди нейрофизиологов, исследующих реальные механизмы, работающие в живом организме [12, 24, 25].

Следует отметить, что в отличие от многочисленных в последнее время рассуждений о «голографическом мозге», «голографической вселенной» и т.п., зачастую основанных лишь на поверхностном знакомстве с внешними эффектами изобразительной голографии, сам Прибрам подчеркивал: «... должно, однако, учитываться предостережение о том, что речь идет о математике голографии и функции мозга, по которой она должна сравниваться и проверяться, а не об оптических голограммах или компьютерной реализации голографии». Для оптика суть важны также и вышеприведенные аналогии, но в любом случае следует иметь в виду, что никакая технология, в том числе, и голография, не может сегодня претендовать на всеохватность решения задачи ИИ, но лишь на частичное решение отдельных, вполне конкретных задач с учетом особенностей, вытекающих из различий материальных носителей БИ и ИИ [13].

Литература к Лекции 1.

1. *Кузнецов О.П.*, Неклассические парадигмы в ИИ // Известия АН, Сер. Теория и системы управления, 1995, №5, с.3-23
2. *Кузнецов О.П., Тарасов В.Б., Аверкин А.Н, Вагин В.Н.*, Круглый стол «Парадигмы искусственного интеллекта»//Новости искусственного интеллекта, 1998, №3, с.140-161.
3. *Kosko B.* Neural Networks and Fuzzy Systems: A dynamical systems approach to machine intelligence - Englewood Cliffs, Prentice Hall int.ed.,1992.

4. *Финн В.К.* Искусственный интеллект: идейная база и основной продукт. Труды Девятой Национальной Конференции по Искусственному Интеллекту, 28.09 – 02.10.04, Тверь, М.Физматлит, 2004, Т.1, с.11-20.
5. *Cowan N.* Attention and Memory. An Integrated Framework. - Oxford Psychology Series, №26, Oxford, 1995
6. *Арбиб М.* Метафорический мозг - М.Мир, 1976
7. *Прибрам К.* Языки мозга - М.Прогресс, 1975.
8. *Марр Д.* Зрение. Информационный подход к изучению представления и обработки зрительных образов. (пер. с англ.) - М. Радио и связь, 1987.
9. *Васильев В.Н., Беспалов В.Г.* Информационные технологии, оптический компьютер и фотонные кристаллы // Сб. «Проблемы когерентной и нелинейной оптики», СПб, ИТМО.2000, с.88-109.
10. *Кузнецов О.П., Шипилина Л.Б.* Псевдооптические нейронные сети – полная прямолинейная модель и методы расчета ее поведения // Изв. АН, Серия: Теория и системы управления, 2000, №5, с.168-176.
11. *Заде Л.*, Понятие лингвистической переменной и его применение к принятию приближенных решений // Математика. Новое в зарубежной науке. М.: Мир, 1976. в.3
12. *Судаков К.В.*, Голографический принцип системной организации процессов жизнедеятельности. - Успехи физиол.наук. 1997. 28, С.3-32.
13. *Прибрам К.*, «Нелокальность и локализация: голографическая гипотеза о функционировании мозга в процессе восприятия и памяти», // В сб. «Синергетика и психология». Вып.1. "Методологические вопросы", Издательство МГСУ "Союз", 1997.
14. *Пригожин И.Р., Стенгерс И.*, «Время, хаос, квант», Пер. с англ. М.Наука, 1989.
15. *Гибсон Дж.*, «Экологический подход к зрительному восприятию», Пер. с англ., М., Прогресс, 1988.
16. *Курейчик В.М.*, Генетические алгоритмы, Новости искусственного интеллекта, 1998, 3, с.14-63.
17. *Romanov S.*, Whether Neuron Network can be Intellectual System? //VIII International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Book of Abstracts, 2002, p.80.
18. *Zadeh L.A., Nikravesh M., Aminzadeh F.*, Computing with Words and the BISC Decision Support System // Proceedings of IV Int. Conference on Soft Computing and Measurements, 2002, v.1, pp.18-21.
19. *Нариньяни А.С.* НЕ-факторы 2004// Труды Девятой Национальной Конференции по Искусственному Интеллекту, 28.09 – 02.10.04, Тверь, М.Физматлит,2004,Т.1, с.420-432.
20. *Гергей Т.* Когнитивные системы – потребность информационного общества и вызов компьютерным наукам // Труды Девятой Национальной Конференции по Искусственному Интеллекту, 28.09 – 02.10.04, Тверь, М.Физматлит, 2004, Т.1, с.3-10.

21. Кузнецов О.П. Когнитивная семантика и искусственный интеллект // Искусственный интеллект и принятие решений. 2012. №4. С. 32-42.
22. Спрингер С., Дейч.Г. Левый мозг, правый мозг. М.: Мир, 1983. 256 с.
23. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р. Модели динамики нейронной активности при обработке информации мозгом – итоги «десятилетия» // Успехи Физических наук. 2002. Т. 172. №10. С. 1189-1214.
24. Судаков К.В. Голографический принцип системной организации поведения: развитие представлений И.П. Павлова о ядре и рассеянных элементах анализаторов // Российский медико – биологический вестник им. академика И.П. Павлова. 1999. № 1 – 2. С. 8 – 14.
25. Александров Ю.И., Брушлинский А.В., Судаков К.В., Умрюхин Е.А. Системные аспекты психической деятельности – М.: Эдиториал УРСС, 1999.

Лекция 2. Базовые математические операции и их реализация оптическими методами

Комплексное представление гармонического сигнала

Из школьного курса математики известно, что гармонический сигнал может быть представлен в комплексной форме, что удобно с точки зрения формального описания. Это представление базируется на следующих соотношениях:

$$\begin{aligned}\cos \varphi &= 0.5(\exp j\varphi + \exp -j\varphi) \\ j \sin \varphi &= 0.5(\exp j\varphi - \exp -j\varphi).\end{aligned}$$

Отсюда следует

$$\exp j\varphi = \cos \varphi + j \sin \varphi \Rightarrow 2A \cos \varphi = A \operatorname{Re}(\exp j\varphi).$$

Плоская монохроматическая волна в точке (x,y,z) в момент времени t может быть представлена в виде

$$s(x, y, z, t) = v(x, y, z)u(t),$$

где

$$\begin{aligned}v(x, y, z) &= E \exp(j\mathbf{k}\mathbf{r}) \exp(-j\varphi) = E \exp[j(k_x x + k_y y + k_z z)] \exp(-j\varphi); \\ u(t) &= \exp(-j\omega t)\end{aligned}$$

Здесь E – амплитуда колебаний электрического поля \mathbf{k} – волновой вектор, его модуль $k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \frac{2\pi}{\lambda} = \frac{\omega n}{c}$, где λ – длина волны, ω – круговая частота, n – показатель преломления среды, φ – начальная фаза волны при $t=0$. При рассмотрении процессов взаимодействия когерентных

полей зависимость от времени не важна и для описания поля достаточно члена

$$E \exp[j(k_x x + k_y y + k_z z)].$$

Напомним, что *комплексная форма записи волновых полей физического смысла не имеет!*

Преобразование Фурье

Операция преобразования Фурье функции $f(x)$.

Пусть функция $f(x)$ удовлетворяет следующим условиям:

1. Функция $f(x)$ абсолютно интегрируема;
2. Функция $f(x)$ непрерывна или имеет конечное число разрывов первого рода и конечное число максимумов и минимумов в любых конечных пределах;
3. Функция $f(x)$ не имеет разрывов второго рода.

Тогда фурье-образ функции $f(x)$ определяется посредством следующего преобразования (преобразование Фурье функции $f(x)$):

$$F(\nu) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi\nu x) dx, \quad (2.1)$$

где ν – пространственная частота, ω – круговая пространственная частота $\omega = 2\pi\nu$.

Обратное преобразование Фурье связывает фурье-образ с его прообразом (т.е. с самой функцией $f(x)$):

$$f(x) = \int_{-\infty}^{\infty} F(\nu) \exp(j2\pi\nu x) d\nu. \quad (2.2)$$

Так как $F(\nu) d\nu = F(\omega) d\omega$, то $F(\nu) = F(\omega) \frac{d\omega}{d\nu} = F(\omega) \frac{2\pi d\nu}{d\nu} = 2\pi F(\omega)$.

Отсюда:

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \exp(-j\omega x) dx \quad (2.3)$$

Отметим, что формально прямое преобразование Фурье отличается от обратного только знаком под экспонентой. Выбор этого знака не принципиален и является предметом договоренности. Мы выбрали для прямого преобразования Фурье знак минус.

В оптике операция Фурье-преобразования реализуется тонкой положительной линзой – при помещении в ее переднюю фокальную плоскость транспаранта с изображением, освещаемого плоским волновым фронтом, в задней фокальной плоскости формируется Фурье-образ (пространственно-частотный спектр) этого изображения как поля комплексных амплитуд в плоскости изображения.

Прежде, чем перейти к обсуждению физического смысла преобразования Фурье, осуществляемого тонкой линзой, рассмотрим ряд вопросов.

Фурье-преобразование функции двух независимых переменных.

Выше мы определили фурье-преобразование для одномерной функции, но в оптике волновые поля (их комплексные амплитуды), описываются функциями двух координат. Фурье-образ функции двух независимых переменных, удовлетворяющей вышеприведенным условиям, может быть определен аналогично (2.1), т.е. как функция двух независимых переменных ν и μ , называемых пространственными частотами:

$$F(\nu, \mu) = \int_{-\infty}^{\infty} f(x, y) \exp(-j2\pi(\nu x + \mu y)) dx dy. \quad (2.4)$$

Обратное преобразование определяется образом, аналогичным (2.2).

Нетрудно видеть, что свойства Фурье-преобразования, определенные для функции одной переменной, справедливы и для функции двух (и более) переменных. В то же время в общем случае функция $F(\nu, \mu)$ не сводима к ее сечениям по отдельным координатам, т.е. замена исследования функции двух переменных отдельным исследованием двух функций, каждая из которых – функция одной переменной, неправомерно. Однако существуют частные случаи, когда упрощение аналитических выражений возможно – это функции с разделяющимися переменными.

Функцией с разделяющимися переменными (в определенной системе координат) называется такая функция двух независимых переменных, которая может быть представлена в виде произведения двух функций, каждая из которых зависит только от одной независимой переменной. Например, в Декартовой системе координат для функции с разделяющимися переменными $f(x, y) = f(x)f(y)$ нетрудно убедиться, что и Фурье-образ (спектр) также будет функцией с разделяющимися переменными $F(\nu, \mu) = F(\nu)F(\mu)$.

Функцией с осевой симметрией в цилиндрической системе координат называется такая функция с разделяющимися переменными $f(\rho, \varphi) = f(\rho)f(\varphi)$, значения которой определяются только радиус-вектором ρ .

Далее мы рассмотрим наиболее важные примеры Фурье-преобразований в предположении, что переменные в функции разделяются.

Прямоугольный импульс – здесь следует рассмотреть два случая:

1. Применительно к оптике прямоугольная функция от двух координат может рассматриваться как прямоугольная апертура. Таким образом, в данном случае преобразование Фурье описывает дифракцию Фраунгофера на прямоугольной апертуре. Нетрудно видеть, что прямоугольная апертура есть функция с разделяющимися в Декартовых координатах переменными т.е.:

$$A(x, y) = \begin{cases} 1 & \text{при } |x| \leq L_x, |y| \leq L_y \\ 0 & \text{при } |x| > L_x, |y| > L_y \end{cases} = A(x) \cdot A(y)$$

$$\begin{aligned} FA(x) &= \int_{-L}^L A(x) \exp(-j\omega_x x) dx = \frac{1}{-j\omega_x} \exp - j\omega_x x \Big|_{-L_x}^{L_x} = \\ &= \frac{\exp j\omega_x L_x - \exp(-j\omega_x L_x)}{j\omega_x} = \frac{2\text{Sin}(\omega_x L_x)}{\omega_x} = 2L_x \text{Sinc}(\omega_x L_x) \end{aligned}$$

И аналогичное выражение для фурье-образа по второй координате $F(A(y))$, т.е. картина дифракции Фраунгофера на прямоугольной апертуре описывается выражением:

$$F(A(x, y)) = 2L_x \text{Sinc}(\omega_x L_x) 2L_y \text{Sinc}(\omega_y L_y). \quad (2.5)$$

2. Дифракция Фраунгофера на круглой апертуре. Здесь мы имеем дело с функцией, обладающей осевой симметрией в цилиндрической системе координат, т.е. $A(\rho, \phi) = A(\rho)$. Для такой функции преобразование Фурье в цилиндрической системе координат будет иметь вид:

$$\begin{aligned} F(A(\rho, \phi)) &= \int_0^{2\pi} d\phi \int_0^{\infty} A(\rho, \phi) \exp(-j2\pi k \rho \cos(\phi - \theta)) \rho d\rho = \\ &= 2\pi \int_0^{\infty} A(\rho) \rho d\rho \frac{1}{2\pi} \int_0^{2\pi} \exp(-j2\pi k \rho \cos(\phi - \theta)) d\phi \end{aligned}$$

Поскольку внутренний интеграл в данном выражении есть функция Бесселя нулевого порядка

$$J_0(2\pi k \rho) = \frac{1}{2\pi} \int_0^{2\pi} \exp - j2\pi k \rho \cos(\phi - \theta) d\phi,$$

то искомым фурье-образ не зависит от фазового угла θ , но зависит лишь от пространственной частоты k , т.е. имеем окончательное выражение

$$F(k) = 2\pi \int_0^{\infty} A(\rho) J_0(2\pi k \rho) \rho d\rho \quad (2.6)$$

Этот вид двумерного преобразования Фурье для осесимметричных функций известен также как преобразование Фурье-Бесселя или преобразование Ганкеля нулевого порядка. Если функция апертуры может быть определена следующим образом:

$$A(\rho) = \begin{cases} 1 & \text{if } \rho \leq \rho_0 \\ 0 & \text{if } \rho > \rho_0 \end{cases},$$

то получим фурье-образ осесимметричной функции в следующем виде:

$$F(\kappa) = 2\pi \int_0^{r_0} J_0(2\pi\kappa r) r dr = \pi r_0^2 \frac{2J_1(2\pi\kappa r_0)}{2\pi\kappa r_0}.$$

Необходимо отметить, что, несмотря на внешнюю схожесть функций (2.5) и (2.6), замена одной функции другой, что, к сожалению, иногда встречается, ведет к существенным ошибкам в результатах. Для примера укажем лишь, что радиус первого нуля в картине дифракции Фраунгофера для прямоугольной апертуры $r_0 = \lambda f/D$, а для круглой – $r_0 = 1.22 \lambda f/D$.

Гармоническое колебание, промодулированное прямоугольным импульсом – дифракционная решетка с синусоидальным профилем штриха в прямоугольной апертуре. Здесь мы воспользуемся свойством делимости переменных и для упрощения выкладок рассмотрим одномерный случай:

$$\begin{aligned} \cos(\omega_0 x) &= \frac{1}{2} (\exp j\omega_0 x + \exp -j\omega_0 x) \\ F(A(x) \cos(\omega_0 x)) &= \frac{1}{2} F(A(x) \exp j\omega_0 x) + \frac{1}{2} F(A(x) \exp -j\omega_0 x) = \quad (2.7) \\ &= LSinc(L(\omega + \omega_0)) + LSinc(L(\omega - \omega_0)) \end{aligned}$$

Выражение (2.7) полезно с точки зрения перехода к следующему этапу – рассмотрению физического смысла преобразования Фурье, осуществляемого линзой. Смысл фурье-разложения заключается в представлении любой непрерывной на заданном интервале функции, удовлетворяющей условиям Дирихле, в виде суммы гармонических колебаний различных частот. Применительно к оптической обработке информации это означает, что любое изображение, которое может быть аналитически представлено в виде функции пропускания (для транспаранта) или яркости (для объекта) от пространственных координат, может быть представлено в виде суммы дифракционных решеток с синусоидальным профилем штриха различных периодов и дифракционной эффективности. Поскольку угол дифракции на решетке зависит от ее периода (частоты решетки), то волна, дифрагировавшая на каждой такой решетке, в соответствии с (2.7) фокусируется в фокальной плоскости линзы в точках, отстоящих от главной оптической оси линзы на расстояние $r = \lambda f \nu$, где f – фокусное расстояние линзы, ν – пространственная частота решетки, измеряемая в обратных миллиметрах (мм^{-1}). Сумма этих дифракционных максимумов, образованных дифрагировавшими на каждой элементарной решетке волнами и описываемых (2.7), и образует фурье-спектр (фурье-образ) изображения.

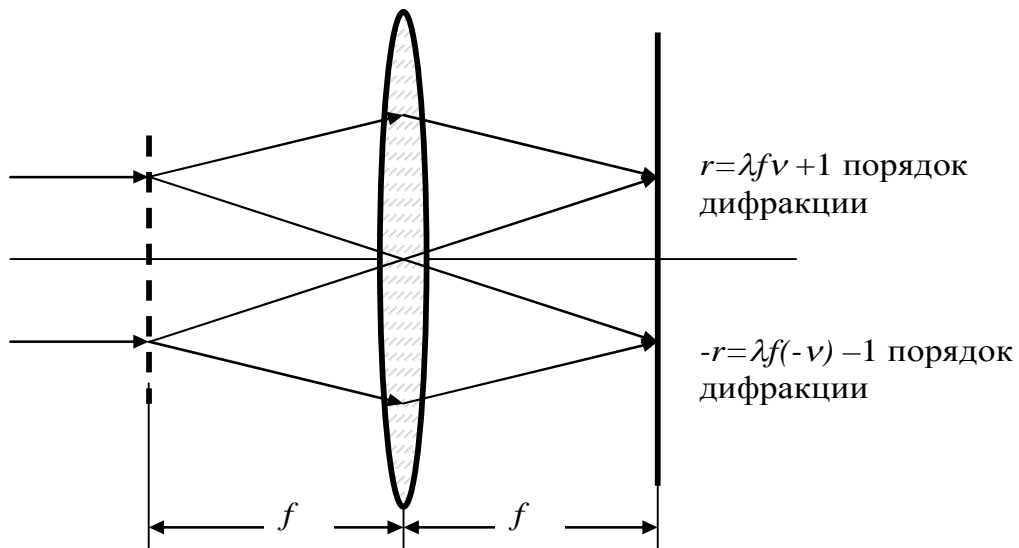


Рис.2.1. Дифракция света на одной решетке – механизм, лежащий в основе формирования Фурье-спектра изображения как суммы решеток

Свойства преобразования Фурье.

Свойство симметрии фурье-образа:

$$F(f^*(x)) = \int_{-\infty}^{\infty} f^*(x) \exp(-j\omega x) dx = \left(\int_{-\infty}^{\infty} f(x) \exp(-j(-\omega)) x dx \right)^* = F^*(-v)$$

Изменение масштаба аргумента функции (Рис.2.2):

$$F(f(ax)) = \int_{-\infty}^{\infty} f(ax) \exp(-j\omega x) dx = \frac{1}{a} \int_{-\infty}^{\infty} f(\xi) \exp\left(-j \frac{\xi \omega}{a}\right) d\xi = \frac{1}{a} F\left(\frac{v}{a}\right) \quad (2.8)$$

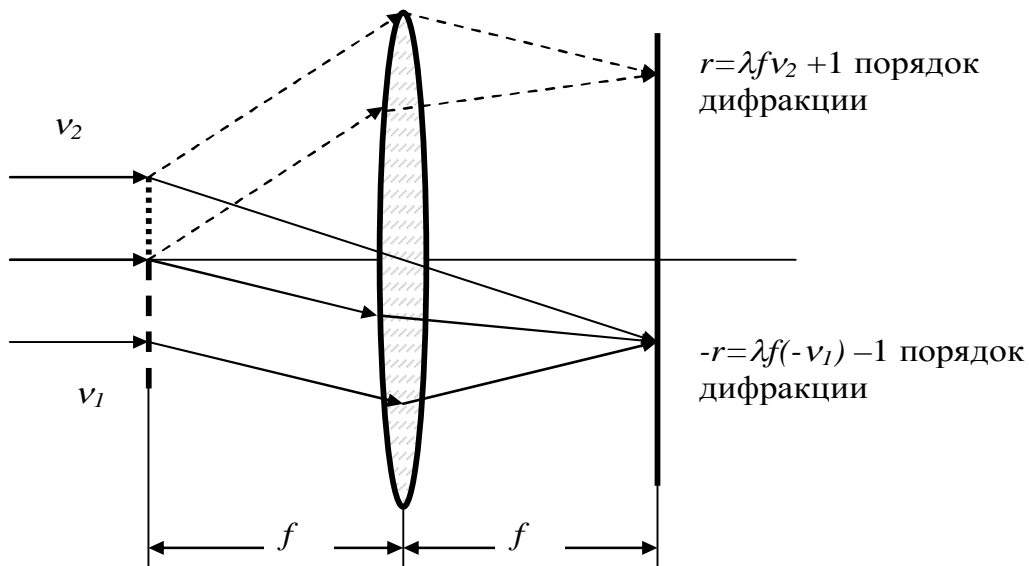


Рис.2.2. Изменение масштаба аргумента функции

Сдвиг аргумента функции (Рис.2.3):

$$\begin{aligned}
 F(f(x-x_0)) &= \int_{-\infty}^{\infty} f(x-x_0) \exp(-j\omega x) dx = \\
 &= \int_{-\infty}^{\infty} f(x_1) \exp(-j\omega(x_1+x_0)) dx_1 = F(f(x_1)) \exp(-j\omega x_0) = \\
 &= F(v) \exp(-j\omega x_0)
 \end{aligned}
 \tag{2.9}$$

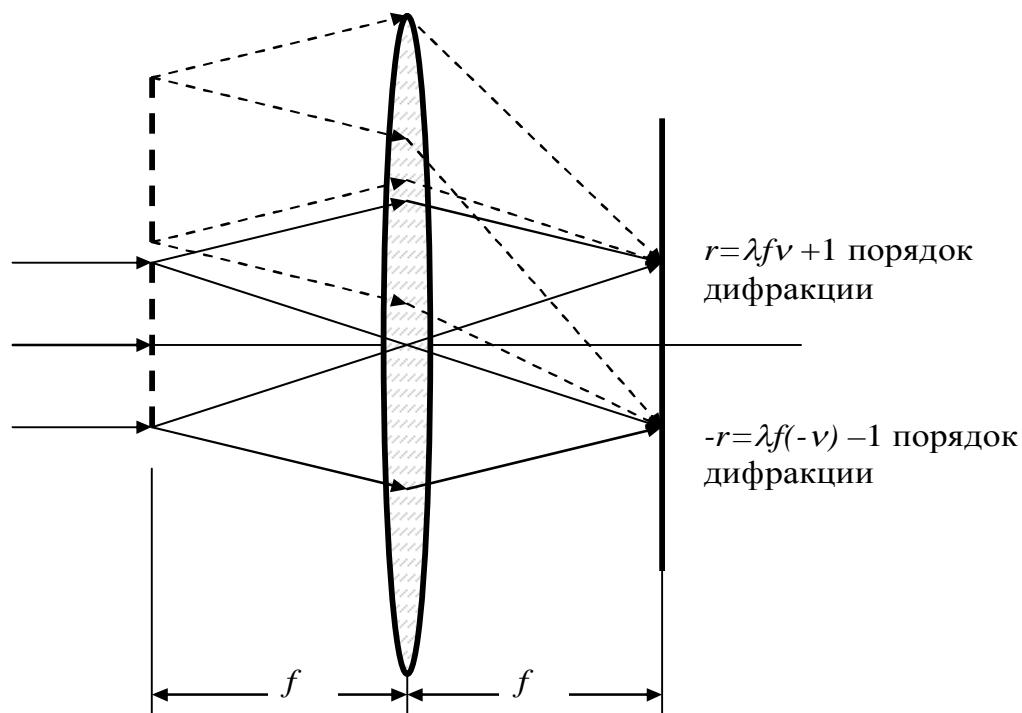


Рис.2.3. Сдвиг аргумента функции

Смещение Фурье-образа по шкале частот:

$$\begin{aligned}
 F^{-1}(F(v+v_0)) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(v+v_0) \exp(j\omega x) d\omega = \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(v_1) \exp(j(\omega_1 - \omega_0)x) d\omega_1 = \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(v_1) \exp(j\omega_1 x) \exp(-j\omega_0 x) d\omega_1 = f(x) \exp(-j\omega_0 x)
 \end{aligned}
 \tag{2.10}$$

Интегральные операции, реализуемые через преобразование Фурье

Операция свертки двух функций определяется следующим образом:

$$S(\Delta) = f_1(x) * f_2(x) = \int_{-\infty}^{+\infty} f_1(x) f_2(\Delta - x) dx . \quad (2.11)$$

Операция корреляции двух функций изначально получила развитие в теории случайных процессов. Функция ковариации двух случайных процессов (или функций) $X(t)$ и $Y(t)$ - это второй центральный (или центрированный) момент, определяемый для комплексных функций следующим образом [4]:

$$R_{XY}(t + \Delta, t) = \left\langle \left(X(t + \Delta) - \langle X(t + \Delta) \rangle \right) \left(Y(t) - \langle Y(t) \rangle \right)^* \right\rangle ,$$

где угловые скобки означают усреднение по ансамблю (набору реализаций). Термин "центральный" означает нормировку на первый момент - математическое ожидание процесса.

Нетрудно видеть, что если случайные процессы стационарны в широком смысле, т.е. в смысле независимости двух первых моментов от начала отсчета, то корреляционная функция зависит только от сдвига одной функции относительно другой Δ .

Если $X(t) = Y(t)$, то имеем автоковариационную функцию процесса (АКФ). В этом случае при $t = s$ получаем дисперсию процесса:

$$R_{XX}(t + \Delta, t) \Big|_{\Delta = 0} = \left\langle \left(X(t) - \langle X(t) \rangle \right) \left(X(t) - \langle X(t) \rangle \right) \right\rangle = \sigma_X^2 .$$

Функция корреляции определяется:

$$K_{XY}(t + \Delta, t) = \left\langle X(t + \Delta) Y^*(t) \right\rangle .$$

Если процесс центрирован, то корреляция и ковариация совпадают.

Отметим, что в литературе встречаются и другие определения функции корреляции, когда сдвиг придается аргументу второго процесса, либо меняется его знак, либо комплексно сопряженным берется первый член. Эти варианты обусловлены теми или иными соображениями при выборе направления осей в пространствах функций и корреляций. Мы будем использовать приведенное выше определение и, как увидим ниже, оно адекватно той физической реальности, что является предметом нашего рассмотрения в этом курсе.

Коэффициент корреляции определяется как корреляционная функция, нормированная на её максимальное значение, т.е. значение в точке $\Delta = 0$:

$$\rho(\Delta) = \frac{K(\Delta)}{K(0)} .$$

Корреляционная функция показывает статистическую связь значений процесса, отстоящих друг от друга на расстояние Δ . Вводится понятие

радиуса (или длины) корреляции r_α , который определяется как значение Δ , на котором коэффициент корреляции убывает до заданного значения:

$$r_\alpha : \rho(r_\alpha) = \alpha; \alpha \in [0,1].$$

Понятие корреляции нашло применение и при обработке сигналов и изображений, являющихся детерминированными функциями. Строго говоря, применение аппарата корреляционного анализа к детерминированным функциям должно быть основано на трактовке детерминированных функций как реализаций стационарного случайного процесса в предположении выполнения условия эргодичности (идентичности усреднения по ансамблю усреднению по реализации). Для этого достаточно выполнения условия Слуцкого, которое правомочно при удовлетворении справедливого для реальных сигналов и полей условия:

$$\lim_{\Delta \rightarrow \infty} K(\Delta) = 0.$$

Отметим, что даже в случае детерминированных функций, которые не могут трактоваться как реализации случайных процессов, понятие корреляции правомочно и практически значимо в силу того, что:

- показывает внутреннюю связность их элементов (внутреннюю коррелированность функций), являющуюся атрибутом, отличающим информацию от дельта-коррелированного шума (для последнего $r_\alpha = 0; \alpha \in [0,1]$), т.е. даже ближайшие значения полностью независимы);
- служит мерой сходства функций, поскольку при $\Delta = 0$ корреляционная функция представляет собой скалярное произведение:

$$K_{XY}(0) = \langle (X(t))(Y^*(t)) \rangle.$$

В рамках курса нас будет интересовать именно последний аспект корреляционной теории, т.е. её применение к обработке детерминированных функций, и мы будем использовать следующее определение функции взаимной корреляции двух функций:

$$K_{f_1 f_2}(\Delta) = f_1(x) \otimes f_2(x) = \int_{-\infty}^{+\infty} f_1(x + \Delta) f_2^*(x) dx. \quad (2.12)$$

Из (2.12) следует, что в общем случае функция корреляции – комплексная функция.

Свойства операций свертки и корреляции

Коммутативность свертки:

$$f_1(x) * f_2(x) = \int_{-\infty}^{+\infty} f_1(x) f_2(\Delta - x) dx = \int_{-\infty}^{+\infty} f_1(\Delta - x) f_2(x) dx = f_2(x) * f_1(x).$$

Некоммутативность операции корреляции:

$$K_{f_1 f_2}(\Delta) = f_1(x) \otimes f_2(x) = \int_{-\infty}^{+\infty} f_1(x + \Delta) f_2^*(x) dx \Big|_{x + \Delta = z} =$$

$$\int_{-\infty}^{+\infty} f_1(z) f_2^*(z - \Delta) dz = \left(\int_{-\infty}^{+\infty} f_2(z - \Delta) f_1^*(z) dz \right)^* = K_{f_2 f_1}^*(-\Delta) \quad (2.13)$$

Отсюда видно, что если $f_1(x) = f_2(x)$, то функция корреляции, являющаяся в данном случае функцией автокорреляции или автокорреляционной функцией (АКФ), для действительных функций обладает свойством четности $K_{f_1 f_1}(-\Delta) = K_{f_1 f_1}(\Delta)$. Для комплексных функций имеем $K_{f_1 f_1}(-\Delta) = K_{f_1 f_1}^*(\Delta)$.

Заметим, что если согласно (2.12) корреляция вычисляется перемножением значений функций $f_1(x)$ и $f_2(x)$ при сдвиге первой на величину Δ , то из (2.13) следует правомочность сдвига второй функции $f_1(x)$ относительно второй, но в противоположном направлении. Свойство некоммутативности операции корреляции понадобится нам в дальнейшем, при переходе к рассмотрению голографических методов вычисления корреляции и реализующих их устройств – голографических корреляторов.

Прямая теорема свертки – фурье-образ свертки двух функций равен произведению их фурье-образов:

$$F(f_1(x) * f_2(x)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x) f_2(\Delta - x) dx \exp(-j\omega\Delta) d\Delta \Big|_{\Delta - x = z} =$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x) f_2(z) dx \exp(-j\omega x) \exp(-j\omega z) dz =$$

$$= \int_{-\infty}^{+\infty} f_1(x) \exp(-j\omega x) dx \int_{-\infty}^{+\infty} f_2(z) \exp(-j\omega z) dz = F_1(\nu) F_2(\nu)$$

Прямая теорема корреляции – Фурье-образ функции корреляции двух функций равен произведению Фурье-образа одной функции на комплексно сопряженный Фурье-образ другой функции:

$$\begin{aligned}
F(f_1(x) \otimes f_2(x)) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x + \Delta) f_2^*(x) dx \exp(-j\omega\Delta) d\Delta \Big|_{x + \Delta = z} = \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(z) f_2^*(x) dx \exp(-j\omega(z-x)) dz = \\
&= \int_{-\infty}^{+\infty} f_1(z) \exp(-j\omega z) dz \int_{-\infty}^{+\infty} f_2^*(x) \exp(j\omega x) dx = \\
&= \int_{-\infty}^{+\infty} f_1(z) \exp(-j\omega z) dz \left(\int_{-\infty}^{+\infty} f_2(x) \exp(-j\omega x) dx \right)^* = F_1(\nu) F_2^*(\nu)
\end{aligned} \tag{2.14}$$

Обратим внимание, что если мы используем другое, часто встречающееся в литературе представление интеграла корреляции

$$\int_{-\infty}^{+\infty} f_1(x) f_2^*(x - \Delta) dx,$$

то мы должны не забыть учесть свойство некоммутативности данной операции (2.13) в плане правильного знака сдвига Δ . Тогда получим корректный результат

$$\begin{aligned}
F(f_1(x) \otimes f_2(x)) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x) f_2^*(x - \Delta) dx \exp(-j\omega\Delta) d(-\Delta) \Big|_{x - \Delta = z} = \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_1(x) f_2(z) dx \exp(-j\omega(x-z)) dz = \\
&= \int_{-\infty}^{+\infty} f_1(x) \exp(-j\omega x) dx \int_{-\infty}^{+\infty} f_2^*(z) \exp(j\omega z) dz = \\
&= \int_{-\infty}^{+\infty} f_1(x) \exp(-j\omega x) dx \left(\int_{-\infty}^{+\infty} f_2(z) \exp(-j\omega z) dz \right)^* = F_1(\nu) F_2^*(\nu) = K_{f_1 f_2}(\Delta)
\end{aligned}$$

Теорема Котельникова о дискретном представлении непрерывных сигналов с ограниченным спектром

Владимир Александрович Котельников (24.08 (06.09) 1908, Казань, Российская империя – 11.02.2005, Москва, РФ). Доказал теорему в 1933 г. [5]

Теорема Котельникова имеет чрезвычайно важное практическое значение, так как лежит в основе методов дискретизации непрерывных сигналов и, тем самым, в основе всех современных методов компьютерной обработки и передачи по каналам связи аудио и видео информации.

Применительно к задаче ИИ она представляет интерес в силу того, что информация из внешнего мира поступает на сенсоры в виде непрерывных функций (или нам так кажется), а как сами сенсоры, так и обрабатывающие ее структуры (нейронные структуры), дискретны.

В англоязычной литературе используется термин «теорема Найквиста-Шеннона», в последнее время и в России многие стали забывать об отечественном приоритете и называть теорему Котельникова теоремой Найквиста-Шеннона. Гарри Найквист в своей работе 1928 года «Certain topics in telegraph transmission theory» [6] исследовал вопрос необходимой полосы частот линии связи для передачи импульсного сигнала и показал, что верхняя частота полосы пропускания должна быть не менее удвоенной частоты следования импульсов. Эта частота сегодня известна как «частота Найквиста». Но вопрос восстановления аналогового сигнала из дискретного Найквист не рассматривал – он изучал только вопрос передачи дискретного сигнала. Клод Шеннон доказал теорему о восстановлении аналогового сигнала из дискретного, аналогичную теореме Котельникова, в 1949 году [7], т.е. на 16 лет позже Котельникова.

Отметим, что в математике вопрос восстановления аналогового сигнала из дискретного рассматривался и до Котельникова, Найквиста и Шеннона. В частности, еще в конце XIX века, в 1897 году великий Феликс Эдуард Жустин Эмиль Борель сформулировал первую часть теоремы [7].

Пусть имеется сигнал $f(x)$ со спектром $F(\nu)$, ограниченным максимальной частотой ν_{max} :

$$f(x) = \int_{-\nu_{max}}^{\nu_{max}} F(\nu) \exp j 2\pi \nu x d\nu \quad (2.15)$$

Поскольку пределы интегрирования ограничены пределами $-\nu_{max}$ ν_{max} , то можно продолжить интегрируемую функцию $F(\nu)$ (спектр) вне этих пределов посредством ее периодического повторения с периодом $2\nu_{max}$. Тогда эту новую функцию (спектр) можно представить рядом Фурье:

$$F(\nu) = \frac{1}{2} \sum_{k=-\infty}^{k=+\infty} \dot{F}_k \exp j 2\pi \frac{k\nu}{2\nu_{max}}, \quad (2.16)$$

где

$$\dot{F}_k = \frac{2}{2\nu_{max}} \int_{-\nu_{max}}^{\nu_{max}} F(\nu) \exp - j 2\pi \frac{k\nu}{2\nu_{max}} d\nu \quad (2.17)$$

Подставив выражение (2.16) в формулу (2.15), получим:

$$f(x) = \int_{-\nu_{max}}^{\nu_{max}} \left(\frac{1}{2} \sum_{k=-\infty}^{k=+\infty} \dot{F}_k \exp j 2\pi \frac{k\nu}{2\nu_{max}} \right) \exp j 2\pi \nu x d\nu.$$

Изменяя порядок интегрирования и суммирования получим:

$$f(x) = \frac{1}{2} \sum_{k=-\infty}^{k=+\infty} \dot{F}_k \int_{-v_{\max}}^{v_{\max}} \exp\left(j2\pi v \left(x + \frac{k}{2v_{\max}}\right)\right) dv.$$

Нетрудно видеть, что при $x = -k/2v_{\max}$:

$$f\left(-\frac{k}{2v_{\max}}\right) = \int_{-v_{\max}}^{v_{\max}} F(v) \exp - j2\pi \frac{kv}{2v_{\max}} dv,$$

откуда следует:

$$\dot{F}_k = \frac{2}{2v_{\max}} f\left(-\frac{k}{2v_{\max}}\right),$$

а интеграл

$$\begin{aligned} & \int_{-v_{\max}}^{v_{\max}} \exp\left(j2\pi v \left(x + \frac{k}{2v_{\max}}\right)\right) dv = \\ & = \frac{\exp\left(j2\pi v_{\max} \left[x + \frac{k}{2v_{\max}}\right]\right) - \exp\left(-j2\pi v_{\max} \left[x + \frac{k}{2v_{\max}}\right]\right)}{j2\pi \left[x + \frac{k}{2v_{\max}}\right]} = \\ & = 2v_{\max} \text{Sinc}\left(j2\pi v_{\max} \left[x + \frac{k}{2v_{\max}}\right]\right) \end{aligned}$$

Тогда получим выражение для исходной функции как сумму отсчетов:

$$f(x) = \sum_{k=-\infty}^{k=+\infty} f\left(-\frac{k}{2v_{\max}}\right) \text{Sinc}\left(2\pi v_{\max} \left[x + \frac{k}{2v_{\max}}\right]\right).$$

Поскольку k принимает значения от $-\infty$ до $+\infty$, то знаки при k можно поменять местами, а величину $1/2v_{\max}$ обозначить Δx , тогда получим окончательное выражение для дискретного представления непрерывной функции с ограниченным спектром:

$$f(x) = \sum_{k=-\infty}^{k=+\infty} f(k\Delta x) \text{Sinc}\left(2\pi v_{\max} [x - k\Delta x]\right).$$

т.е. непрерывная функция со спектром, ограниченным максимальной частотой v_{\max} может быть представлена посредством отсчетов, отстоящих друг от друга на расстояние $\Delta x = 1/2v_{\max}$.

А поскольку любой реальный процесс или поле всегда ограничены (во времени или пространстве), то и описывающая и функция также ограничена, следовательно, число отсчетов также конечно.

К вопросу биологической обоснованности использования преобразования Фурье в искусственных нейронных сетях

Основоположник фурье-анализа рассматривал предложенный им метод как сугубо абстрактный, формальный прием, не имеющий никакого отношения к реальности. Выше мы убедились, что тут он ошибался, т.к. по крайней мере в оптике преобразование Фурье реально выполняется реальным оптическим элементом – положительной линзой. Но имеет ли разложение по Фурье какое-либо отношение к биологическим интеллектуальным системам? Оказывается, что самое непосредственное.

Как показали результаты экспериментальных исследований зрительной системы, проведенных группой под руководством В.Д. Глезера и, затем, Ю.Е. Шелепина из института нейрофизиологии РАН им. И.П.Павлова, в зрительной системе высших животных и человека реализуется преобразование Фурье [9 - 13]. Более того, имеет место не однократное, но двукратное преобразование Фурье, дающее в итоге реализацию согласованной фильтрации. Гипотеза о согласованной фильтрации в зрительной системе была высказана Н.Н.Красильниковым еще в 1957 году [14] и затем экспериментально подтверждена в нейрофизиологических исследованиях [15-18].

Согласно [13], биологические структуры зрительной системы реализуют два типа согласованных фильтров:

1. нелинейные, согласованные с деталями воспринимаемого изображения – информация с нелинейных фильтров в норме поступает в левое полушарие головного мозга;
2. линейные, согласующие число периодов в весовой функции с размером изображения и создающие абстрактное, т.е. без деталей, фурье-описание изображения – информация с этих фильтров поступает в правое полушарие.

Взаимодействие обоих полушарий дает в итоге фильтр, согласованный с полным детальным описанием изображения, что было подтверждено психофизиологическими экспериментами с переносом кодов из одного полушария в другое [13].

Литература к Лекции 2

1. *Кольер Р., Беркхарт Л., Лин Л.*, "Оптическая голография", М.Мир, 1973, 686с.
2. *Василенко Г.И., Цибулькин Л.М.*, «Голографические распознающие устройства», М., Радио и связь, 1985. – 312с.
3. *Мирошников М.М.*, «Теоретические основы оптико-электронных приборов», Л., Машиностроение, 1977.-600с.

4. Яглом А.М. Корреляционная теория стационарных случайных функций / Л.: Гидрометеиздат, 1981. - 280 С.
5. Котельников В. А. О пропускной способности «эфира» и проволоки в электросвязи // *Успехи физических наук* : — 2006. — № 7. — С. 762-770.
6. H. Nyquist Certain topics in telegraph transmission theory. Trans. AIEE, vol. 47, pp. 617—644, Apr. 1928.
7. C. E. Shannon Communication in the presence of noise. Proc. Institute of Radio Engineers. Vol. 37. No. 1. P. 10—21. Jan. 1949.
8. Erik Meijering A Chronology of Interpolation From Ancient Astronomy to Modern Signal and Image Processing, Proc. IEEE, 90, 2002. DOI:10.1109/5.993400.
9. Глезер В.Д., Иванов В.А., Щербач Т.А. Ответ рецептивных полей зрительной коры кошки на сложные стимулы // *Физиологический журнал СССР*. 1972. Т. 58. № 3.
10. Глезер В.Д., Иванов В.А., Щербач Т.А. Исследование рецептивных полей нейронов зрительной коры кошки как фильтров пространственных // *Физиологический журнал СССР*. 1973. Т. 59. № 2.
11. Глезер В.Д., Дудник К.Н., Куперман А. М., Лушина Л. И., Невская А.А., Подвитин Н.Ф., Праздникова И.В. Зрительное опознание и его нейрофизиологические механизмы. — Л.: Наука, 1975.
12. Глезер В.Д. Зрение и мышление. СПб.: Наука, 1993. 341 С.
13. Глезер В.Д. О роли пространственно-частотного анализа, примитивов и межполушарной асимметрии в опознании зрительных образов // *Физиология человека*. 2000. Т. 26. №5. С. 145 – 150.
14. Н.Н. Красильников, О.И. Красильникова, Ю.Е. Шелепин Экспериментальное исследование согласованной пространственной фильтрации в зрительной системе человека при наблюдении чисто хроматических изображений // *Оптический журнал*. 1999. Т.66. №10. с.22 – 25.
15. Н.Н. Красильников, Ю.Е. Шелепин Функциональная модель зрения // *Оптический журнал*. 1997. Т.64. №2. с.72 – 82.
16. В.Т.Шуваев Согласованная фильтрация – одна из возможных функций базальных ганглиев? // *Оптический журнал*. 1999. Т.66. №10., с.41 – 45.
17. С.В. Алексеенко Архитектура связей в зрительной коре и зрительное опознавание // *Оптический журнал*. 1999. Т.66. №10. с.52 – 54.
18. Ю.Д. Кропотов Система селекции действий мозга как фильтр высокого уровня // *Оптический журнал*. 1999. Т.66. №10. с.55 – 57.

Лекция 3. Распознавание образов. Оптические методы распознавания

В настоящее время можно выделить две крайние точки зрения на роль и место задачи распознавания в проблематике ИИ. Согласно одной из них, распознавание образов – центральная, исторически первая и до сих пор не решенная задача ИИ. Согласно противоположной точке зрения ИИ и распознавание – два совершенно различных научных направления. В рамках того определения ИИ, что мы приняли на первой лекции, представляется, что нам ближе первая точка зрения в силу того, что распознавание – первый и необходимый шаг во всякой интеллектуальной деятельности. Прежде, чем информацию обработать, ее надо распознать. Не распознанная информация субъективно информацией и не является, но лишь шумом. Поэтому важнейший этап интеллектуальной деятельности – *восприятие* – может быть определен как процесс субъективизации информации, что предполагает в том числе, и распознавание.

Корреляционный алгоритм в задаче распознавания. Предположим, что нам надо сравнить два изображения и ответить на вопрос об их идентичности. Если изображения достаточно сложные, с большим количеством деталей и не обладают ярко выраженной семантикой, то простейший способ решить задачу – попытаться совместить изображения, например, наложив их одно на другое и рассматривая на просвет. Если изображения идентичны, то без труда удастся найти положение их полного совпадения. Критерий совпадения очень прост – интегральное пропускание двух наложенных и совмещенных изображений будет максимальным, что хорошо заметно глазом. Совмещение по этому критерию выполняется интуитивно, так как при небольшом сдвиге одного из изображений относительно другого интегральное пропускание резко уменьшается. Если изображения не идентичны, то положения с максимальным пропусканием найти не удастся.

Нетрудно убедиться, что описанный алгоритм описывается выражением (2.10), определяющим функцию взаимной корреляции двух функций. Оптические устройства, вычисляющие корреляционную функцию (2.10) вышеописанным методом, известны под названием некогерентных оптических корреляторов. Очевидно, что такой метод вычисления корреляционной функции «в лоб» трудоемок и неэффективен – в каждый момент времени вычисляется корреляционный функционал для текущего значения сдвига и для вычисления всей функции корреляции необходимо просканировать весь диапазон значений сдвига. Решение проблемы было предложено в 1963 г. сотрудником Радиолокационной лаборатории Мичиганского университета (США) Ван дер Люгтом и основано на использовании техники голографии Фурье.

Рассмотрим этот метод подробнее, начав с аналитического описания 4f-схемы фурье-голографии, приведенной на рис.3.1.

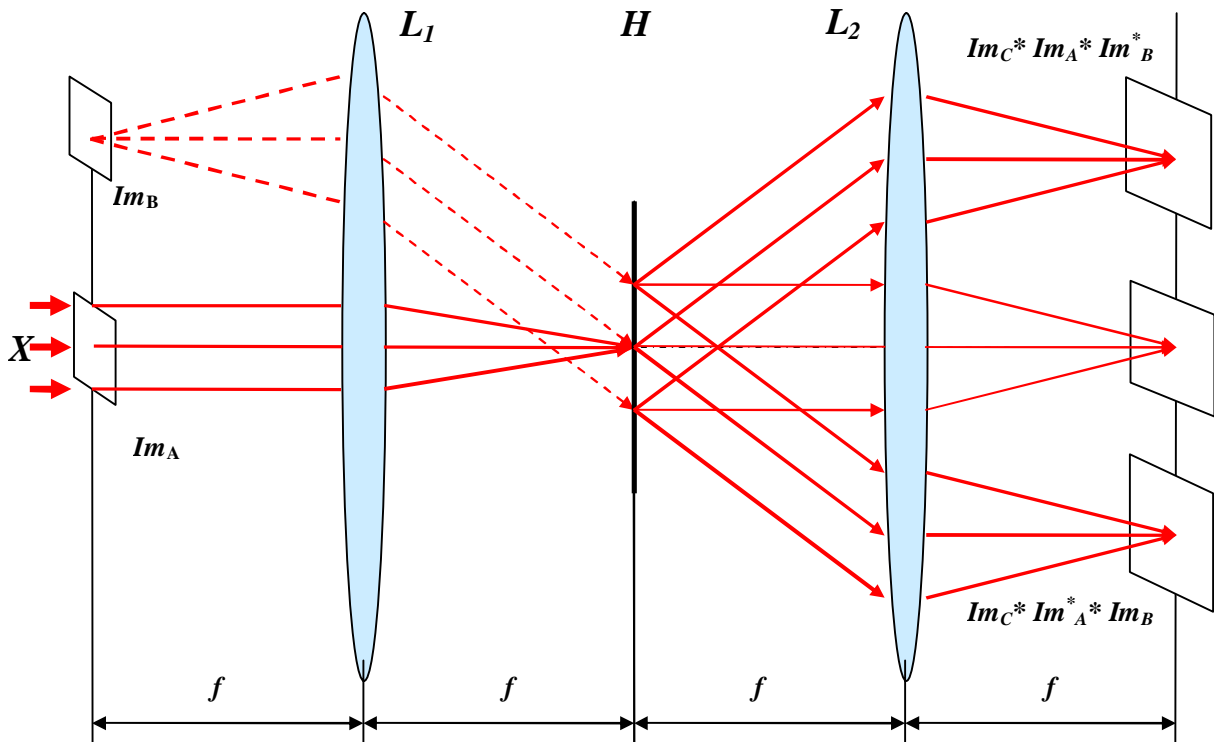


Рис.3.1. 4f-схема голографии Фурье

Уравнение 4f-схемы голографии Фурье. Пусть во входной плоскости **In** рассматриваемой схемы рис.3.1. помещены два транспаранта, функции амплитудного пропускания которых описываются $Im_A(x,y)$ и $Im_B(x,y)$. Далее для того, чтобы не загромождать выражения, координаты опустим. Плоскость **In** суть передняя фокальная плоскость первой Фурье-преобразующей линзы **L1**. Транспаранты освещаются плоским волновым фронтом X . Для простоты выражений примем, что амплитуда X равна 1. Тогда в задней фокальной плоскости линзы **L1** формируется поле:

$$F(Im_A) + F(Im_B).$$

При помещении в эту плоскость регистрирующей среды **H**, на ней после экспонирования и надлежащей обработки (проявления) будет записано:

$$\eta((F(Im_A) + F(Im_B))(F(Im_A) + F(Im_B))^*),$$

где η – оператор регистрирующей среды (в нашем случае – голографической), а астериск означает комплексное сопряжение. При помещении проявленной голограммы в положение, соответствующее положению голографической регистрирующей среды при записи голограммы, и помещении в плоскость **In** вместо изображения Im_A (эталонного) изображения Im_C , называемого объектным, сразу за голограммой получим поле, характеризующееся следующим распределением комплексных амплитуд:

$$F(Im_C)\eta\left(\left(F(Im_A)+F(Im_B)\right)\left(F(Im_A)+F(Im_B)\right)^*\right)=$$

$$F(Im_C)\eta\left(F(Im_A)F(Im_A)^*\right)+F(Im_C)\eta\left(F(Im_A)F(Im_B)^*\right)+$$

$$+F(Im_C)\eta\left(F(Im_B)F(Im_A)^*\right)+F(Im_C)\eta\left(F(Im_B)F(Im_B)^*\right)$$

Члены, входящие в это выражение, описывают компоненты дифрагировавшей на голограмме волны, распространяющиеся в разных направлениях, а именно: волны, описываемые членами

$$F(Im_C)\eta\left(F(Im_A)F(Im_A)^*\right)$$

и

$$F(Im_C)\eta\left(F(Im_B)F(Im_B)^*\right)$$

распространяются по направлению распространения волны $F(Im_C)$ и формируют нулевой порядок дифракции; волна, описываемая членом

$$F(Im_C)\eta\left(F(Im_A)^*F(Im_B)\right)$$

распространяется в направлении распространения волны $F(Im_B)$ при записи и формирует +1 порядок дифракции; а волна, описываемая членом

$$F(Im_C)\eta\left(F(Im_A)F(Im_B)^*\right)$$

распространяется в направлении, комплексно сопряженном с направлением распространения волны $F(Im_B)$ при записи голограммы и формирует -1 порядок дифракции.

Соответственно, в фокальной плоскости второй Фурье-преобразующей линзы L_2 формируются три области – три порядка дифракции: волны

$$F\left(F(Im_C)\eta\left(F(Im_A)F(Im_A)^*\right)\right)$$

и

$$F\left(F(Im_C)\eta\left(F(Im_B)F(Im_B)^*\right)\right)$$

формируют нулевой порядок дифракции, который в рамках нашего рассмотрения нас не интересует;

в +1 порядке дифракции формируется распределение амплитуд:

$$F\left(F(Im_C)\eta\left(F(Im_A)^*F(Im_B)\right)\right), \quad (3.1)$$

а в -1 порядке дифракции формируется распределение амплитуд:

$$F\left(F(Im_C)\eta\left(F(Im_A)F(Im_B)^*\right)\right). \quad (3.2)$$

Если оператор голографической регистрирующей среды η линеен, а $Im_B = \delta$, т.е. точечный источник (напомним, что фурье-образ точечного

источника – плоская волна), то к выражениям (3.1) и (3.2) можно применить теоремы корреляции и свертки, соответственно, что даст:

$$F\left(F(Im_C)\eta\left(F(Im_A)^*F(Im_B)\right)\right)=Im_C\otimes Im_A \quad (3.3)$$

$$F\left(F(Im_C)\eta\left(F(Im_A)F(Im_B)^*\right)\right)=Im_C*Im_A \quad , \quad (3.4)$$

где символы \otimes и $*$ обозначают операции корреляции и свертки. Такую 4f-схему фурье-голографии, в которой используется только +1 порядок дифракции называют также коррелятором Ван дер Люгта, а фурье-голограмму – голографическим согласованным фильтром, поскольку передаточная характеристика голограммы согласована со спектром эталонного изображения Im_A .

Коррелятор совместного преобразования. Другой метод вычисления функции взаимной корреляции двух функций в 4f-схеме фурье-голографии может быть реализован, если в выражениях (3.1) и (3.2) принять за δ -функцию не опорное изображение Im_B , а объектное Im_C . В этом случае они примут вид:

$$F\left(F(Im_C)\eta\left(F(Im_A)^*F(Im_B)\right)\right)=Im_B\otimes Im_A \quad , \quad (3.5)$$

$$F\left(F(Im_C)\eta\left(F(Im_B)^*F(Im_A)\right)\right)=Im_A\otimes Im_B \quad , \quad (3.6)$$

т.е. в +1 и -1 порядках дифракции формируются функции взаимной корреляции изображений, с которых записана голограмма. Понятия опорного и предметного пучков в этом случае физического смысла не имеют. Поскольку на голограмме записываются совместно спектры (точнее - картина интерференции спектров) двух изображений, функция корреляции которых вычисляется, то такая схема известна под названием коррелятора совместного преобразования.

Выбор оператора голографической регистрирующей среды и его роль.

Для того, чтобы применить к выражениям (3.1) и (3.2) теоремы корреляции и свертки, мы приняли допущение о линейности оператора голографической регистрирующей среды η , не обсуждая вопрос что именно мы понимаем, когда говорим об операторе η . В большинстве учебников по голографии в качестве такого оператора рассматривается зависимость пропускания регистрирующей среды τ от экспозиции. Однако, в рамках рассматриваемой нами темы удобнее использовать другую зависимость – зависимость дифракционной эффективности η от экспозиции. Обратим внимание, что речь здесь идет не о дифракционной эффективности всей голограммы в целом (интегральной), а о «локальной» дифракционной эффективности, т.е. о дифракционной эффективности фрагмента голограммы, размер которого определяется в соответствии с изученной в предыдущей лекции теоремой Котельникова. Это замечание существенно в первую очередь именно для голограмм Фурье.

Напомним, что голограмма представляет собой зарегистрированную картину интерференции двух пучков – опорного и сигнального. Соответственно, голограмма Фурье – это зарегистрированная картина интерференция фурье-спектра эталонного изображения (сигнальный пучок) и плоского опорного пучка. Поскольку для фурье-спектра характерен существенный перепад амплитуд, достигающий нескольких порядков, то видность интерференционной картины максимальна только в точке равенства амплитуд сигнального и опорного пучков. (Рис.3.2).

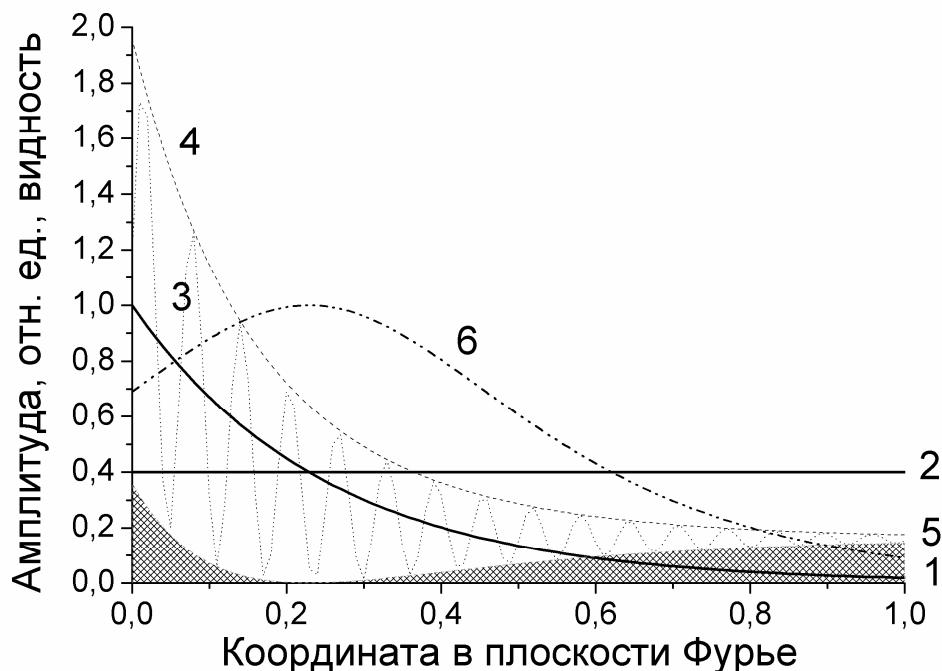


Рис.3.2. Интерференция сигнального и опорного пучков в плоскости Фурье: зависимости от координат в плоскости Фурье: 1, 2 – амплитуд сигнального (1) и опорного (2) пучков, 3 – интерференционных полос, обусловленных углом схождения сигнального и опорного пучков (условно), 4 и 5 – огибающих максимумов и минимумов картины интерференции, 6 – видности интерференционной картины

Во всех остальных точках фурье-плоскости (на частотах, отличных от частоты равенства локальных амплитуд спектра и опорного пучка) видность интерференционной картины будет меньше 1.

Для большинства голографических регистрирующих сред (ГРС) типичная экспозиционная характеристика, т.е. зависимость дифракционной эффективности (ДЭ) от экспозиции, имеет нелинейный характер, пример дан на Рис.3.3, и включает участки прямой и инверсной зависимости ДЭ от экспозиции. Интересующий нас как рабочий, т.е. восходящий (в идеале – квазилинейный) участок экспозиционной характеристики занимает лишь незначительную часть динамического диапазона – обычно 1-2 порядка. Таким образом, вследствие ограниченности динамического диапазона ГРС на фурье-голограмме может быть записан только ограниченный участок спектра. Выбор этого диапазона пространственных частот каждый раз определяется задачей.

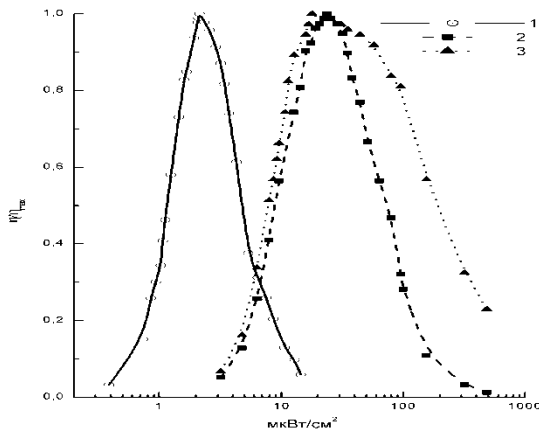


Рис.3.3.а. Зависимости нормированной дифракционной эффективности от экспозиции для структур ХСП-НЖК (стибнит Sb_2S_3) для толщин слоя ЖК: 1- 6.5 мкм, 2 – 10 мкм, 3 – 15 мкм

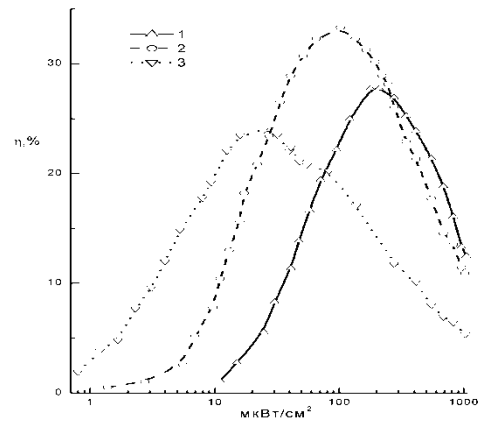


Рис.3.3.б. Зависимости дифракционной эффективности от экспозиции для структур ХСП-НЖК со слоем ИКС-24 для напряжений питания 1 – 10 в., 2 – 15 в., 3 – 25 в. [4]

Если условия записи голограммы подбираются так, чтобы локальная экспозиция в этой точке соответствовала экспозиции верхней границы рабочего диапазона (Рис.3.3), то часть интерференционной картины, находящаяся на рис.3.2 справа от точки равенства амплитуд, т.е. в области более высоких пространственных частот, попадет в рабочий диапазон. Если в рабочем диапазоне экспозиционная характеристика может быть аппроксимирована линейной зависимостью, то этот участок спектра будет записан линейно – фильтр в этом диапазоне согласован со спектром. Верхняя частотная граница согласованного участка определяется нижней точкой рабочего участка на рис.3.3.

Левее точки равенства пучков на рис.3.2 на формирование локальной передаточной характеристики влияет совокупность следующих факторов:

- видность интерференционной картины убывает;
- появляется постоянная немодулированная засветка, возрастающая при удалении от точки равенства пучков;
- зависимость дифракционной эффективности от экспозиции может нелинейно возрастать, быть постоянной или убывать.

В зависимости от удельного веса каждого из этих факторов, Фурье-голограмма может иметь одно из типичных сечений дифракционной эффективности, примеры которых приведены на рис.3.4.

При согласовании голограммы в области высоких пространственных частот, в области пространственных частот ниже частоты равенства сигнального и опорного пучков записывается инверсная голограмма (дифракционная эффективность имеет обратную зависимость от амплитуды спектра), а в центральной части голограммы, т.е. в области нулевых и низких пространственных частот, дифракционная эффективность может быть равна нулю.

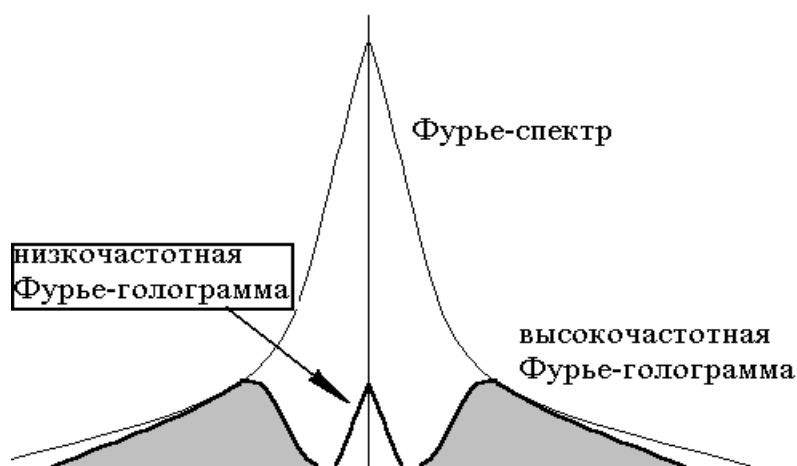


Рис.3.4. Сечения дифракционной эффективности низкочастотной и высокочастотной Фурье-голограмм

Таким образом, в силу нелинейности оператора голографической регистрирующей среды в $+1$ и -1 порядках дифракции $4f$ -схемы фурье-голографии формируются распределения амплитуд, пропорциональные корреляции и свертке не тех изображений Im_A и Im_C , которые предъявлялись при записи и восстановлении голограммы, а изображений, профильтрованных оператором η , т.е. выражения (3.3) и (3.4) более корректно записать в виде, учитывающем дополнительную фильтрацию

$$F(F(Im_C)\eta(F(Im_A)^*F(Im_B))) = Im^{\eta_C} \otimes Im^{\eta_A} \quad (3.7)$$

$$F(F(Im_C)\eta(F(Im_A)F(Im_B)^*)) = Im^{\eta_C} * Im^{\eta_A}. \quad (3.8)$$

Эта дополнительная фильтрация может быть в ряде случаев чрезвычайно полезной, например, при распознавании изображений, подверженных временной изменчивости. В частности, при сравнении изображений природных объектов мелкие детали (листва), которым соответствуют высокие пространственные частоты, подвержены значительным изменениям. Поэтому голограмма, согласованная в области низких пространственных частот, обеспечит большую устойчивость распознавания по сравнению с высокочастотной голограммой.

Является ли голографический коррелятор устройством распознавания?

Голографический коррелятор (Рис.3.1) вычисляет функцию взаимной корреляции эталонного и объектного изображений. Корреляционный функционал удовлетворяет аксиоматическому определению метрики, т.е. может служить мерой близости двух функций. Однако в реальной жизни мы без проблем опознаем человека, которого видели много лет тому назад, в совсем другой обстановке. Без труда опознаем изображение по его части, под другим углом зрения, в другом масштабе и т.п., т.е. мы распознаем *инвариантно* к сдвигу, искажениям, масштабу, фрагментации и дефектам.

Голографический коррелятор обеспечивает инвариантность к сдвигу – при смещении опознаваемого изображения во входной плоскости корреляционное поле синхронно смещается в выходной плоскости. Физически, сдвиговая инвариантность голографического коррелятора суть угловая инвариантность тонкой голограммы. Если голограмма объемная, то угловой инвариантностью она не обладает и теряется инвариантность к сдвигу.

К остальным факторам (масштаб, геометрические искажения, поворот) ни корреляционный алгоритм, ни голографический коррелятор инвариантностью не обладают. Причину помогает понять рис.3.5, на котором изображены дифракционные максимумы эталона (пустые кружки, размер $\lambda f/D$, где D – апертура изображения) и повернутого объектного изображения. Из рис.3.5 также видно, что один из методов уменьшения критичности к искажениям – оптимизация частотного диапазона. На низких частотах перекрытие дифракционных максимумов больше, чем на высоких.

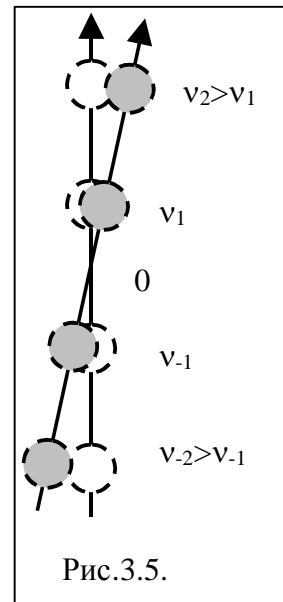


Рис.3.5.

Был предложен целый ряд методов обеспечения инвариантности к тем или иным искажениям. Рассмотрим два из них:

1. Использование иных, нежели Фурье, преобразований, обеспечивающих инвариантность к требуемым искажениям. В качестве примера рассмотрим преобразование Меллина, обеспечивающее инвариантность к масштабу. Преобразование Меллина определяется:

$$M(g(x, y)) = \int g(x, y) x^{-j\omega_x - 1} y^{-j\omega_y - 1} dx dy .$$

Соответственно, для функций $g(x, y)$ и $g(ax, ay)$ имеем:

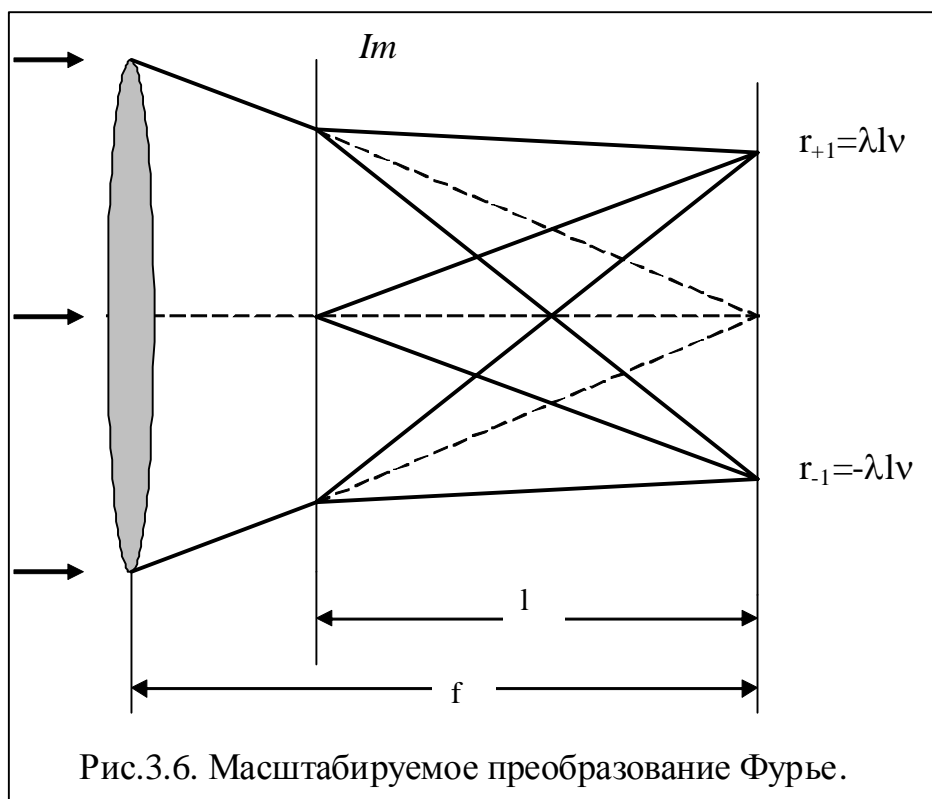
$$M(g(ax, ay)) = a^{-j\omega_x} a^{-j\omega_y} M(g(x, y)),$$

т.е. получаем инвариантность к масштабу. Преобразование Меллина функции $g(x)$ эквивалентно преобразованию Фурье функции $g(\exp(\xi))$. Оптическая реализация преобразования Меллина, таким образом, включает два этапа:

- на первом выполняется логарифмическое преобразование координат функции $g(x, y)$, это преобразование целесообразно выполнять в электронном тракте ввода информации в оптический процессор;
- затем с помощью линзы Фурье-преобразование новой функции.

Однако, нетрудно видеть, что *получив инвариантность к масштабу, мы потеряли инвариантность к сдвигу.*

2. Масштабируемое преобразование Фурье. Если транспарант с изображением поместить не перед фурье-преобразующей линзой, а за ней, как показано на рис.3.6, то, как нетрудно видеть, масштаб фурье-образа будет зависеть от расстояния от транспаранта до фокальной плоскости линзы – угол дифракции остается постоянным, так как он зависит только от длины волны и частоты решетки, а величина r зависит также и от расстояния l . Изменяя это расстояние в соответствии с изменением масштаба изображения, можно подогнать масштаб фурье-преобразования. Однако и в этом случае, *получая инвариантность κ масштабу, мы теряем инвариантность κ сдвигу.*



3. Метод наложенных голограмм.

Интуитивно очевидное решение заключается в том, что поскольку глобальный максимум автокорреляционной функции (ГМ АКФ) при появлении искажений деградирует не скачком, но плавно, то следует записать на одной регистрирующей среде несколько голограмм, подобрав их так, чтобы когда одна голограмма переставала «работать» (ГМ АКФ не детектируется), начинала «работать» другая, перекрыв весь возможный диапазон искажений. Однако следует учитывать, что в то время, как одна голограмма «работает», т.е. формирует корреляционный пик, остальные голограммы формируют в той же точке выходной плоскости значения функций кросс-корреляции. В результате, отношение сигнал-помеха снижается как минимум в N раз, где N – число наложенных голограмм. Второй отрицательный эффект этого метода – все голограммы должны быть «втиснуты» в рабочий участок динамического диапазона регистрирующей среды. В результате дифракционная эффективность

каждой голограммы также снижается в N раз. Таким образом, и этот метод, *давая выигрыш в одном, приводит к потерям в другом* – отношении сигнал-помеха и проблемам с энергетикой. Если постараться учесть все возможные в реальной ситуации искажения, то метод окажется неработоспособен – число потребных голограмм слишком велико.

Можно рассмотреть еще ряд методов, предложенных для обеспечения инвариантного распознавания, и при внимательном изучении окажется, что *всегда выполняется один закон, который можно назвать законом сохранения степеней свободы – выигрывая в одном, неизбежно теряем в другом*. Таким образом, голографический коррелятор представляет из себя высокопроизводительный процессор – вычислитель функции взаимной корреляции определенных на плоскости функций, но для решения задачи инвариантного распознавания сам по себе он еще не пригоден – нужно какое-то «обрамление», какой-то иной подход. То рассмотрение, что мы провели выше, полностью укладывается в рамки компьютерной парадигмы. Таким образом, мы сталкиваемся с упоминавшейся на первой лекции проблемой поиска новой парадигмы.

В заключение лекции отметим, что есть ряд задач, которые голографический коррелятор решает очень эффективно. Например, задача автосопровождения цели. Опознавание и захват цели осуществляется оператором, по его команде записывается голограмма и цель сопровождается коррелятором. При деградации сигнала ниже определенного порога голограмма автоматически перезаписывается и автосопровождение продолжается.

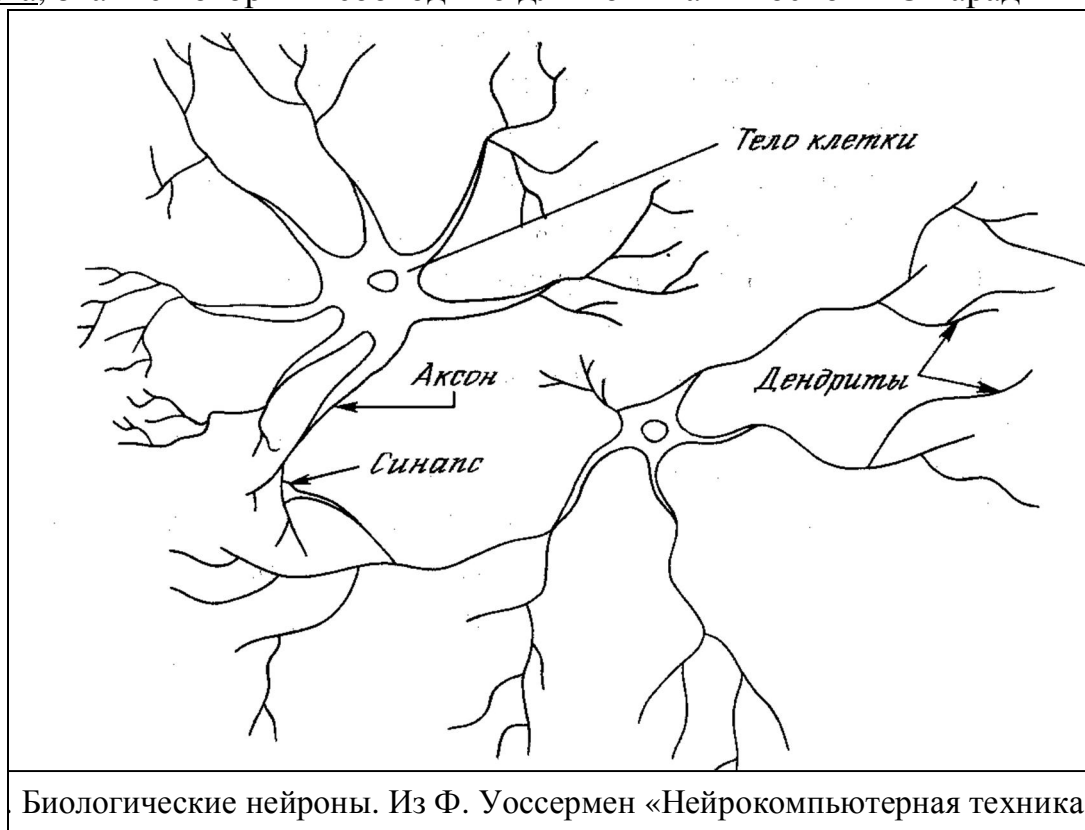
Литература к Лекции 3

1. *Кольер Р., Беркхарт Л., Лин Л.*, "Оптическая голография", М.Мир, 1973, 686с.
2. *Василенко Г.И., Цибулькин Л.М.*, «Голографические распознающие устройства», М., Радио и связь, 1985. – 312с.
3. *Гуревич С.Б., Константинов В.Б., Соколов В.К., Черных Д.Ф.*, «Передача и обработка информации голографическими методами», М., Сов.радио, 1978. –304с.
4. *Амосова Л.П., Плетнева Н.И., Чайка А.Н.* Реверсивная регистрирующая среда типа фотопроводник – жидкий кристалл с выпрямляющим полупроводниковым переходом стибнит – германиевое стекло // Оптический журнал. 2006. Т. 73. № 9. С. 64 – 67.

Лекция 4. Основы парадигмы искусственных нейронных сетей

Нейросетевая парадигма (НС) базируется на той идее, что если создать техническими средствами структуру, подобную по своей конструкции и свойствам биологической структуре мозга, то и работать она будет подобным образом. Поэтому НС парадигма опирается на результаты анатомических и нейрофизиологических исследований мозга.

Нервная система человека состоит из примерно 10^{11} нейронов, связанных в единую сеть посредством более чем 10^{15} связей. О работе отдельного нейрона известно очень много. Но на вопрос о том, каким образом реализуется мышление, нейрофизиология исчерпывающего ответа не дает, поскольку в ней нет конструктивной законченной теории, и, следовательно, и понимания того, как из совокупности химических и физических процессов возникает мышление как качественно новый феномен. Ниже дадим только основные сведения о работе биологического мозга, знание которых необходимо для понимания основ НС парадигмы.



По мнению большинства нейрофизиологов, мышление реализуется объединенными в сеть нервными клетками — нейронами, которых в мозге человека по разным данным от 70 до 100 млрд. Нейроны имеют серый цвет — это те самые «серые клеточки», «серое вещество» мозга.

В последнее время появляется все больше данных об активной роли в мышлении помимо нейронов также и нейроглиальных клеток: астроцитов, олигодендроцитов, микроглиальных клеток. Нейроглиальных клеток на порядок больше, чем нейронов; они образуют нейроглию — ту среду, в которую погружены нейроны.

Нейрофизиологами выявлены разнообразные функции нейроглиальных клеток. Наиболее изучены опорная, трофическая, защитная и барьерная функции. Биологическая защита реализуется благодаря способности микроглиоцитов к фагоцитозу – поеданию бактерий. А астроциты плотно обвивают сосуды головного мозга и являются непреодолимым барьером для многих веществ. Это так называемый гематоэнцефалический барьер, т.е. барьер между кровью и нейронами мозга – это барьерная функция.

Менее изучена роль нейроглии в процессах мышления. Например, роль олигодендроцитов в процессах запоминания новых моторных навыков или влияние астроцитов на функционирование синапсов – синаптическую проводимость сигналов. Об этом поподробнее скажем чуть ниже.

Нейрон имеет много входов (до нескольких сотен) и один выход – *аксон*, который обычно разветвляется на *коллатерали*, заканчивающиеся на разных нейронах. Это эквивалентно тому, что от нейрона отходит много связей, передающих информацию о его возбуждении по многим адресам.

На вход нейрона может приходиться дискретный электрический импульс (возбуждающий или тормозящий) от другого нейрона или сенсора – рецепторной клетки. Амплитуда и полярность передаваемых импульсов постоянны, меняется частота их следования. Мембрана нейрона имеет исходный потенциал. Приходящие на нейрон пакеты импульсов создают возбуждающие или тормозящие потенциалы, которые суммируются (со знаком) на его мембране, изменяя её исходный потенциал – происходит поляризация или деполяризация мембраны. Обычно считают, что на мембране происходит и пространственное, и временное суммирование дополнительных потенциалов. Со временем, если не происходит возбуждения и разряда нейрона, то этот накопленный дополнительный потенциал поляризации или деполяризации его мембраны уменьшается.

Все передающиеся по связям импульсы имеют одинаковую полярность. Импульс становится возбуждающим или тормозящим, т.е. приобретает полярность, проходя через синапс – соединение между входной связью и телом нейрона. Части нейрона, на которых расположены синапсы, называются *дендритами*. Считается, что синапс может изменять свою проводимость, благодаря чему он определяет как знак, так и величину передаваемого потенциала.

Биологический нейрон является дискретным пороговым элементом. По достижении суммарным потенциалом, накопленным на мембране, определенного порога, происходит разряд, формирующий импульс, передающийся по аксону на другие нейроны. Предполагается, что по длине аксона импульс не затухает. Аксоны могут иметь разную толщину и разную длину. Скорость передачи импульса по аксону зависит от толщины волокна и может варьироваться в пределах от 1 до 100 метров в секунду.

После разряда нейрон в течение некоторого времени (около миллисекунды) становится невозбудимым – наступает период абсолютной рефрактерности. Затем следует период длительностью несколько

миллисекунд, в течение которого порог возбудимости нейрона снижается до нормального уровня – это период относительной рефрактерности. Затем – экзальтация, т.е. кратковременное снижение порога возбудимости ниже нормального уровня. Поскольку структура мозга состоит из различных типов нейронов, отличающихся по своим свойствам, то отличаются и периоды рефрактерности и экзальтации у разных нейронов.

Интеллектуальные способности человека определяются не столько числом нейронов, сколько развитостью межнейронных связей. Важность проблемы связности отражена в существовании такого термина, как «*коннекционистская парадигма*» – это синоним НС парадигмы. Связность является и главной проблемой при реализации НС моделей методами электронной техники в силу необходимости локализовать и изолировать каждую связь от всех других. Оптические технологии позволяют решить эту проблему за счет использования явлений дифракции и интерференции, и благодаря распространению оптических волн в свободном пространстве.

Искусственный нейрон. В 1943г. появилась статья МакКалока и Питтса «Логическое исчисление идей, относящихся к нервной деятельности», в которой была предложена модель формального нейрона и, основанные на этой модели формальные логические сети. Основная идея заключалась в попытке описать работу нейронных механизмов мозга, используя аппарат математической логики. Искусственный нейрон представляет собой сумматор, выход которого описывается выражением

$$Out = Nl \sum In ,$$

где Nl – нелинейная активационная функция нейрона, In – входные сигналы (рис.4.2.а).

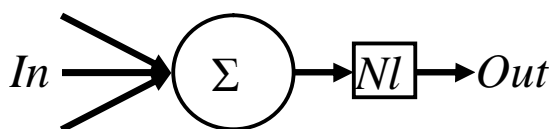


Рис.4.2.а. Нейрон как нелинейный сумматор

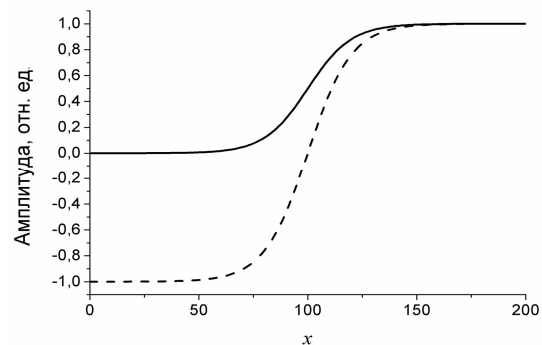


Рис.4.2.б. Нелинейные активационные функции нейрона: сигмоидальная (сплошная) и гиперболический тангенс (штрих)

Нейрон МакКалока и Питтса был бинарным, поскольку техника того времени позволяла легко реализовать такой нейрон посредством реле.

Свойства нейронов и НС МакКалока и Питтса определялись аксиомами:

- Возбуждение нейрона соответствует принципу «все или ничего».
- Время делится на дискретные моменты – такты.
- Возбуждение нейрона в какой-то момент времени происходит, если в предшествующий момент времени возбуждено определенное

фиксированное число синапсов. Это число не зависит ни от предыдущей активности, ни от расположения синапсов на нейроне.

- Возбуждение по связи от одного нейрона к любому другому происходит без задержки (за один такт).
- Синапсы могут быть возбуждающими или тормозящими. Входной сигнал, прошедший через тормозящий синапс, абсолютно исключает возбуждение данного нейрона в рассматриваемый момент времени,
- С течением времени структура сети не изменяется.

Нейронные сети Мак-Каллока и Питтса могут реализовать любую двоичную логическую функцию и, следовательно, любое высказывание логики высказываний. Как показал Клини, сетью Мак-Каллока и Питтса можно реализовать любой конечный автомат.

Модель Мак-Каллока и Питтса вызвала шквал критики как со стороны физиологов и психологов, так и со стороны инженеров. Это обычная судьба большинства пионерских, основополагающих работ. Наряду с критикой, работа стимулировала целый ряд как нейрофизиологических исследований, касающихся свойств отдельного нейрона, и сети в целом, так и исследований в области математической логики и, шире, в области искусственного интеллекта.

Гроссберг (1973) обнаружил, что логистическая или “сигмоидальная” функция $y(x) = \frac{1}{1 + \exp(-x)}$, или подобная S-образная нелинейная характеристика решает поставленную им *дилемму шумового насыщения*.

Сущность проблемы шумового насыщения заключается в том, что нейрон (как и нейронная сеть в целом) должен реагировать только на сигналы, амплитуда которых находится в определенном диапазоне. На сигналы меньшей амплитуды (например, фоновый шум от уличного движения в городе) нейрон реагировать не должен. Эта проблема решается за счет низкого коэффициента усиления передаточной функции нейрона в области малых амплитуд входного сигнала. Аналогично, нейрон должен быть защищен от перевозбуждения при поступлении сигналов слишком высокой амплитуды – в соответствующей области передаточная функция также характеризуется низким коэффициентом усиления и переходом в режим насыщения (Рис.4.2.б).

Гиперболический тангенс по форме сходен с логистической функцией, но принимает также и отрицательные значения, раньше он часто использовался биологами в качестве математической модели активации нервной клетки.

Правило обучения Хэбба. Это правило было сформулировано его автором в вербальной форме: «Если два нейрона одновременно возбуждены, то сила связи между ними увеличивается». Нетрудно видеть, что оно допускает различные варианты математической конкретизации, некоторые из которых мы рассмотрим в этом курсе. Но прежде еще раз вернемся к биологическим нейронным сетям для того, чтобы понять какие

именно механизмы обеспечивают то свойство обучаемости, что представлено правилом Хебба.

В биологической нейронной сети на сегодня можно достаточно условно выделить два метода обеспечения долговременных связей нейронов [2-4]:

1. синаптическая пластичность (СП);
2. миелинизация нервных волокон.

Первый метод работает за счет экспрессии генов в возбужденном нейроне. В результате меняется уровень синтеза рецепторов, медиаторов, первичных и вторичных клеточных посредников, внутриклеточных транспортеров и т.д., что в совокупности ведет к долговременной перестройке функций нейрона.

Экспрессия генов, вызывающих повышение синтеза определенной модификации рецепторов, приводит к увеличению числа этих рецепторов в синапсах и чувствительность синапса меняется – может возрастать, но может и убывать. Экспрессия генов, вызывающих синтез белков цитоскелета, ведет к достройке цитоскелета в области синапса, что влечет лучшее удержание рецепторов в области синапса и увеличение площади синаптического контакта. Это только два механизма долговременной СП.

Однако существует и кратковременная СП, как реакция на которую и возникает долговременная. Кратковременная СП возникает в результате непосредственного изменения свойств постсинаптической мембраны нейрона и её рецепторов как ответ на активность синапсов. Если она не переходит в долговременную, то вся кратковременные изменения в синапсах стираются и информация забывается.

Можно считать, что описанный метод преимущественно обеспечивает первоначальное возникновение связей и, соответственно, гибкость интеллекта, его адаптивность к изменяющимся внешним условиям.

Второй метод ответственен за фиксацию и усиление уже существующих связей, т.е. обеспечивает стабильность. Здесь основанная роль принадлежит олигодендроцитам, составляющим значительную, самую крупную долю нейроглиальных клеток, в среду которых погружены нейроны. Основная функция олигодендроцитов – создание миелиновой оболочки (белого вещества мозга) вокруг нервных волокон [4]. Наличие такой оболочки дает возрастание скорости прохождения импульсов по нервному волокну от 40 до 200 раз. Образно говоря, формируется жесткая структура, защищенная от изменений, происходит «окостенение мозга».

Баланс этих двух реально работающих в биологических структурах механизмов обеспечивает баланс стабильности – пластичности памяти.

Векторно-матричный формализм. Удобно описывать состояния нейронных слоев посредством векторов-строк, а веса связей двух слоев в виде матрицы – произведения векторов $W=A^T B$, где T – символ транспонирования, количество строк m представляет число нейронов слоя A , а количество столбцов n – число нейронов слоя B . Например, $w_{i,j}$ – это вес, связывающий i -ый нейрон слоя A с j -ым нейроном слоя B . Тогда,

вычисление выходного вектора B , компонентами которого являются выходы Out нейронов, сводится к матричному умножению $B = AW$.

Один из примеров формализации правила Хэбба в рамках векторно-матричного формализма

$$w_{ij}(n+1) = w_{ij}(n) + \alpha OUT_i OUT_j,$$

где $w_{ij}(n)$ – значение веса от нейрона i к нейрону j до подстройки, $w_{ij}(n+1)$ – значение веса от нейрона i к нейрону j после подстройки, α – коэффициент скорости обучения, OUT_i – выход нейрона i , OUT_j – выход нейрона j .

Элементарная НС «Звезда Гроссберга “Instar”» дана на рис.4.3.

Здесь и далее кружочками будем обозначать вычислительные нейроны, а квадратиками – сенсорные, не производящие никаких вычислений, а лишь передающие возбуждение на вычислительные нейроны.

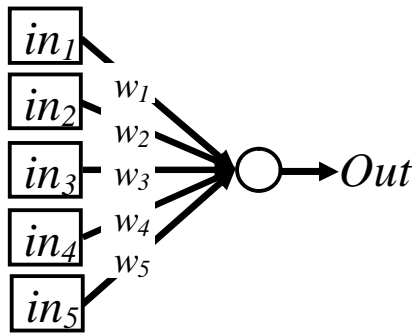


Рис.4.3. Звезда Гроссберга “Instar”

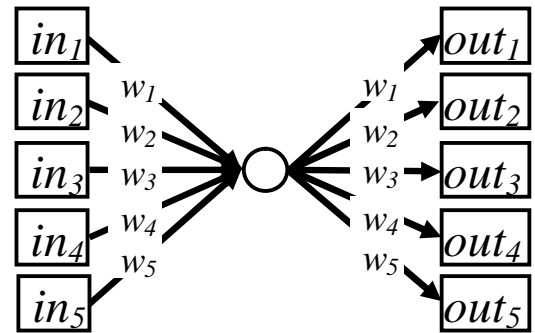


Рис.4.4. Авто-ассоциативная память из двух звезд Гроссберга “Instar” и “Outstar”

Комбинация звезд «Instar – Outstar» позволяет реализовать контекстно-адресуемую или ассоциативную память, толерантную к дефектам, искажениям и шумам опознаваемого образа (Рис.4.4). Если схема симметрична, то память *автоассоциативная*, восстанавливающая входной образ по его искаженной версии. Если звезды различны, то память *гетеро-ассоциативная*, восстанавливающая по искаженному входному образу неискаженный, ассоциированный с ним образ.

Голографический коррелятор как звезда Гроссберга “Instar”. Структура связей, реализуемых в схеме фурье-голографии.

Покажем, что при записи фурье-голограммы формируется и, самое главное – запоминается, структура связей «звезда Гроссберга». Рассмотрим схему записи голограммы Фурье двух точечных источников, представленную на рис. 4.5.1. Обозначим точечные источники символами δ и im_1 , разнесенные друг от друга на расстояние d_1 , и поместим их в переднюю фокальную плоскость Фурье-преобразующей линзы с фокусным расстоянием f . Каждый точечный источник формирует расходящуюся сферическую волну, преобразуемую линзой в волну плоскую. Если источники когерентны, то в задней фокальной плоскости

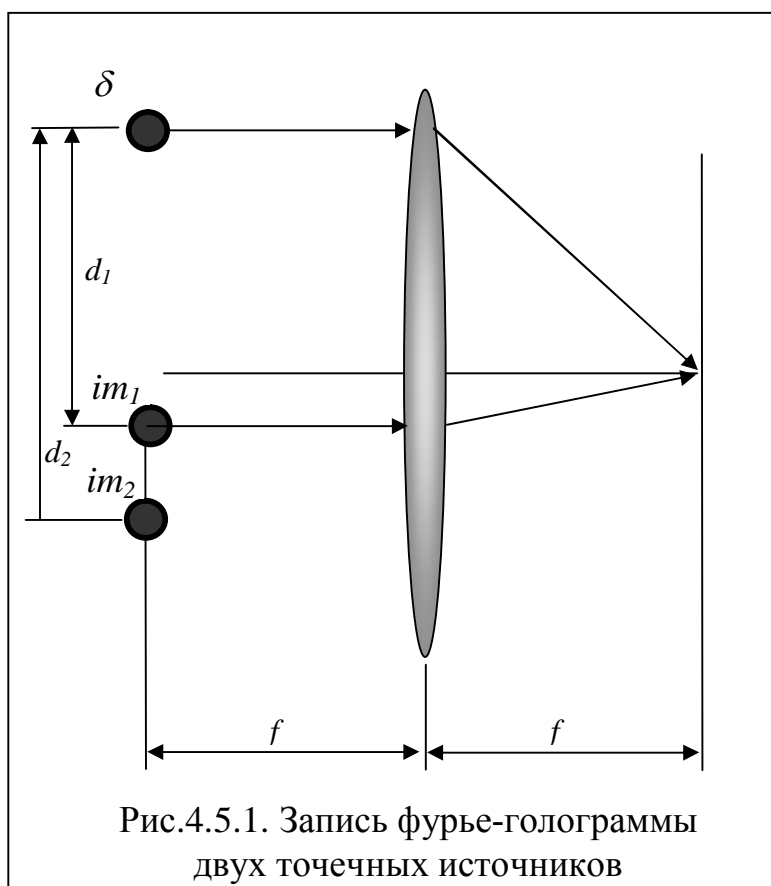
линзы формируется стабильная во времени картина интерференции двух плоских волн – решетка с синусоидальным распределением амплитуды. Частота этой решетки определяется выражением:

$$v_1 = \frac{d_1}{\lambda f},$$

а видность – соотношением амплитуд соответствующих источников:

$$V_1 = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{im_1 + \delta - |im_1 - \delta|}{im_1 + \delta + |im_1 - \delta|}.$$

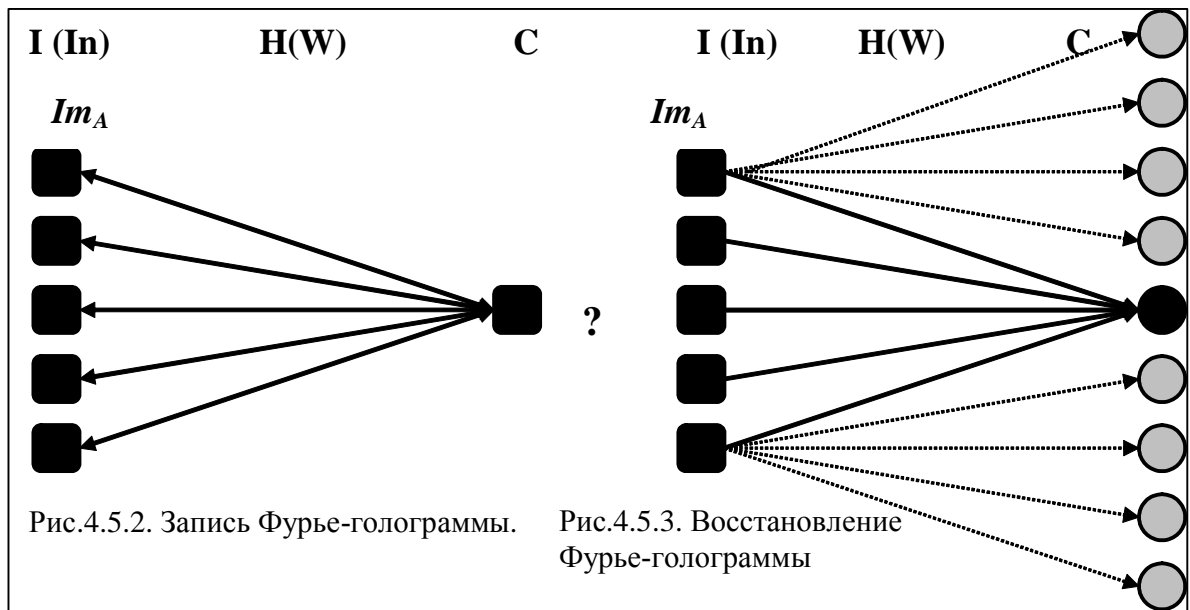
Нетрудно видеть, что видность интерференционной картины является параметром, значение которого соответствует правилу обучения Хэбба – она максимальна (равна 1) тогда, когда амплитуды источников δ и im_1 равны.



Если взять другой точечный источник im_2 на расстоянии d_2 , то картина интерференции будет описываться уже другими значениями параметров v_2 и V_2 . Изображение Im можно представить как набор точечных источников im (пикселей), размер которых определяется в соответствии с теоремой Котельникова. Формируемая в задней фокальной плоскости картина интерференции спектра изображения с плоской опорной волной (фурье-образом точечного опорного источника) представляет собой сумму картин интерференции фурье-образов каждого из пикселей с фурье-образом

опорным источником. Таким образом, каждый точечный источник, из которых состоит изображение, однозначно связан с опорным источником δ параметрами соответствующей интерференционной картины – частота определяется взаимным положением источников, а видность – отношением их амплитуд. Иными словами, при интерференции фурье-образов точечного опорного источника и эталонного изображения, т.е. записи голограммы Фурье, формируется соответствующая модели «звезда Гроссберга» структура связей (Рис.4.5.2), которая реализуется в виде голограммы – записанной интерференционной картины.

Рассматривая голограмму как сумму дифракционных решеток с синусоидальным профилем штриха, очевидно, что дифракционная эффективность каждой такой решетки пропорциональна видности той интерференционной картины, изображением которой эта решетка является. Таким образом, фурье-голограмма реализует правило обучения Хэбба – дифракционная эффективность максимальна у той решетки, которая сформирована при интерференции двух точечных источников равной амплитуды.



Нетрудно убедиться, что фурье-голограмма реализует правило обучения Хэбба – вес связи есть видность интерференционной картины, она максимальна когда уровни возбуждения нейронов равны. При освещении голограммы восстанавливается эта звезда, а δ -нейрон – это нейрон активированный глобальным максимумом АКФ. Но необходимо учитывать, что в силу того, что каждый нейрон как точечный источник формирует расходящуюся сферическую волну, вкупе со свойством угловой инвариантности тонкой голограммы, каждый I-нейрон формирует свою звезду «outstar». В результате, если в эталонном изображении было N_A нейронов, а в объектном N_B , то в корреляционной плоскости активируется N_A+N_B-1 нейронов.

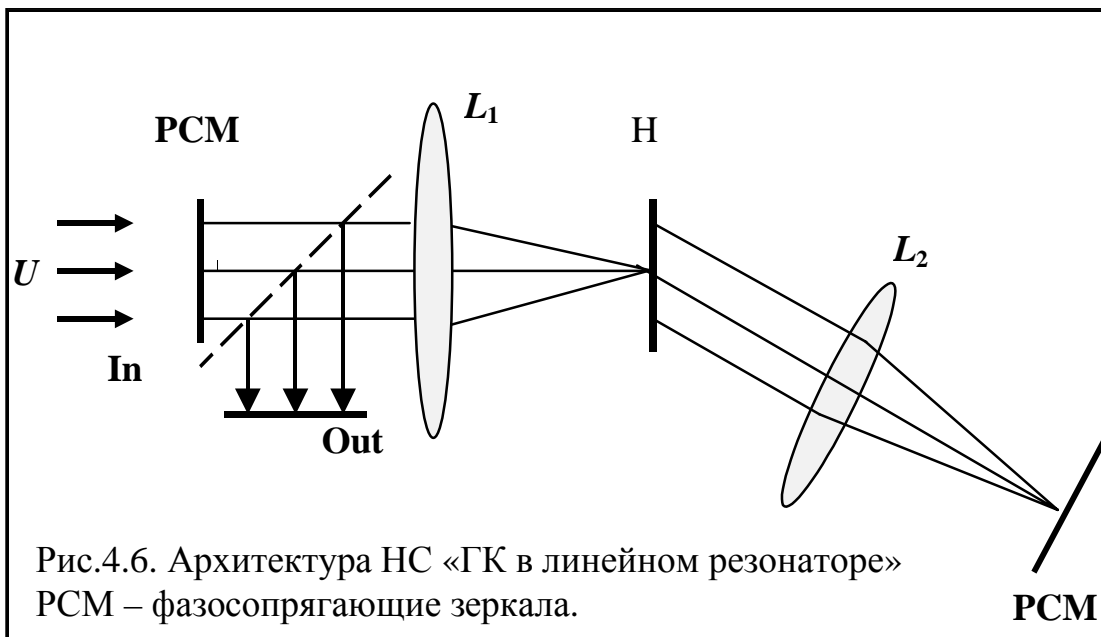


Рис.4.6. Архитектура НС «ГК в линейном резонаторе»
PCM – фазоспрягающие зеркала.

Голографическая авто-ассоциативная память – комбинация звезд «instar-outstar».

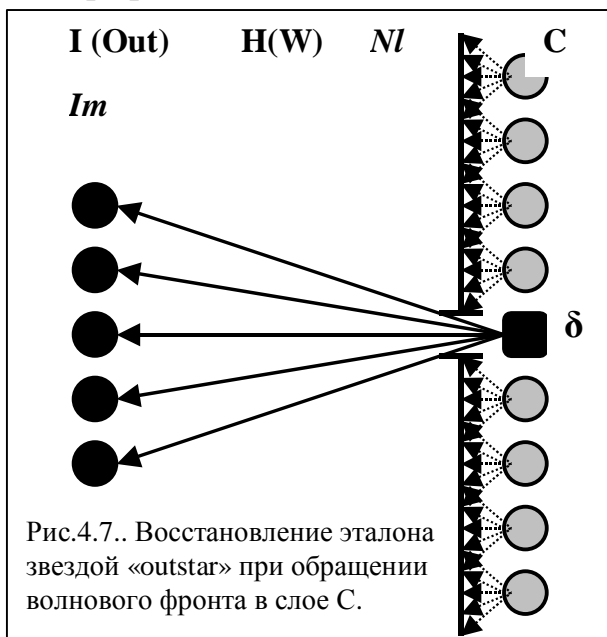


Рис.4.7.. Восстановление эталона звездой «outstar» при обращении волнового фронта в слое С.

Для реализации этой модели в слой С вводится фазоспрягающее зеркало – устройство обращения волнового фронта (ОВФ) и нелинейная функция, режектирующая связи от всех С-нейронов за исключением звезды, восстанавливаемой δ -нейроном – С-нейроном с максимальным уровнем активации. Такая модель выбора δ -нейрона известна под названием WTA – победитель забирает все.

Фазо-сопрягающее зеркало представляет собой динамическую

голограмму, восстанавливаемую пучком, сопряженным опорному (Рис.4.6). Если фазо-сопрягающие зеркала помещаются и во входную и в корреляционную плоскости, то они образуют резонатор, модами которого являются записанные на голограмме образы. При предъявлении какого-либо образа в сети развивается итерационный процесс и, если входной образ подобен одному из записанных на голограмме, то после нескольких итераций возникает резонанс на этой моде – в выходной плоскости формируется устойчивое изображение этого образа. Эта архитектура оптической НС «ГК в линейном резонаторе» известна также как сеть Псалтиса. Структура связей приведена на Рис.4.7. – из всего многообразия связей $C \rightarrow I$ посредством нелинейности в слое С выделяется только одна «звезда Гроссберга», восстанавливающая тот эталон, коэффициент корреляции входного образа с которым максимален.

Нелинейный оператор РСМ Nl подбирается таким образом, чтобы выделить из корреляционного распределения только глобальный максимум автокорреляционной функции (ГМ АКФ)

$$Nl(Im_B \otimes Im_A^\eta) \rightarrow \delta,$$

где δ – дельта-функция, описывающая ГМ АКФ как дифракционно-ограниченный точечный источник. В результате такой процедуры, в обратном ходе лучей $C \rightarrow H \rightarrow Out$, ГМ АКФ как точечный опорный источник восстанавливает в Out изображение Im_A^η

$$Nl(Im_B \otimes Im_A) \ast Im_A^{\eta\ast} = \delta \ast Im_A^{\eta\ast} = Im_A^{\eta\ast},$$

где астериск – символ комплексного сопряжения, а \ast – символ операции свертки. Отметим, что само эталонное изображение Im_A , с которого была записана голограмма H , может быть восстановлено только при линейном операторе η , что для большинства реальных изображений физически невыполнимо в силу фундаментального свойства ограниченности динамического диапазона регистрирующих сред.

Такая голографическая авто-ассоциативная память была разработана, например, на фирме «Boeing» для конструкторско-технологического архива [8]. Общая идея в том, что при разработке нового изделия на вход подается эскиз новой детали или узла и система извлекает из памяти наиболее близкую из ранее разработанных деталей. Тем самым, существенно экономятся время и силы при разработке новой продукции.

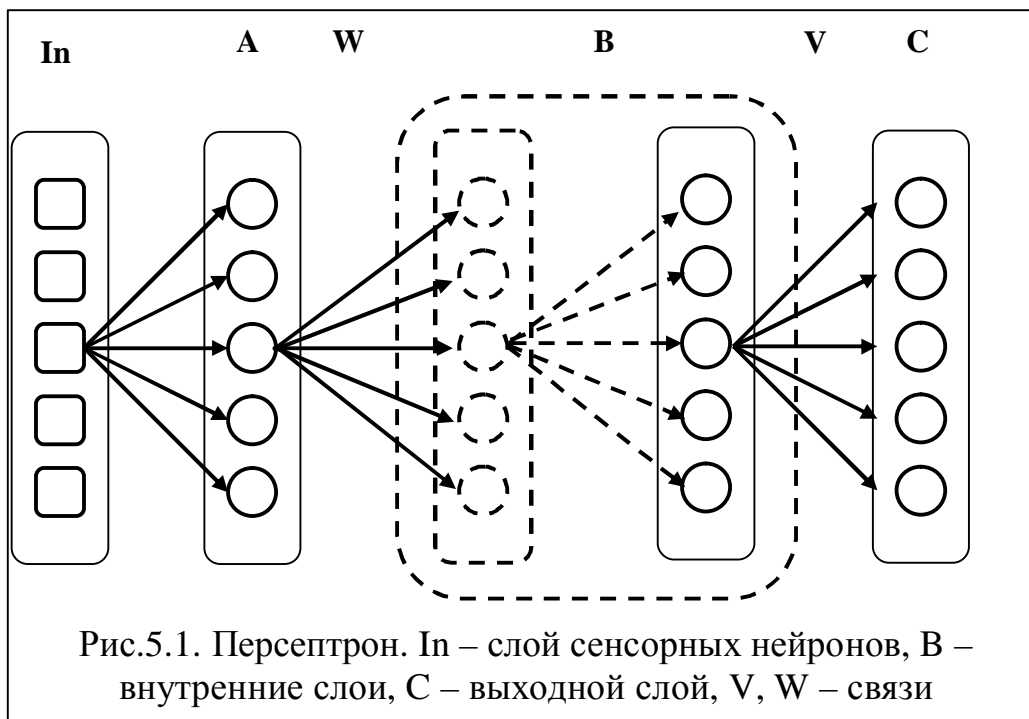
Авторы благодарят А.В. Муллина за ценную помощь в подготовке изложения современных результатов, проясняющих важную и не до конца на сегодня еще исследованную роль нейроглии в процессах формирования межнейронных связей, обеспечении их устойчивости и пластичности.

Литература к Лекции 4.

1. Шамис А. Л., «Нейронные сети и моделирование мышления», Новости Искусственного Интеллекта, 2006, №3, с.26
2. Муллин А.В. Частное сообщение. Март 2017.
3. Xu T, et al. Rapid formation and selective stabilization of synapses for enduring motor memories // Nature. 2009;462:915–919.
4. Xiao L. et. al. Rapid production of new oligodendrocytes is required in the earliest stages of motor skill learning // Nat Neurosci. 2016 Sep; 19(9): 1210–1217.
5. Кольер Р., Беркхарт Л., Лин Л., "Оптическая голография", М.Мир, 1973, 686с.
6. Василенко Г.И., Цибулькин Л.М., «Голографические распознающие устройства», М., Радио и связь, 1985. – 312с.
7. Owechko Y., "Nonlinear holographic associative memories", //IEEE Journal of Quantum Electronics, 1989, v.25, N3, p.619-634.
8. D.C.Wunsh II et.al., Photorefractive adaptive resonance neural network // Applied Optics, 1993, v.32, №8, p.1399-1407.

Лекция 5. Сети однонаправленного распространения

Элементарная НС «звезда Гроссберга» ассоциирует пару векторов, один из которых имеет размерность 1 (в оптике вектор размерностью 1 представляет собой дифракционно ограниченный точечный источник).



Такая структура связей хорошо подходит для тех задач, где требуется простой отклик на входное воздействие, например, задачи распознавания. Но есть задачи, где необходимо сопоставить входному образу другой, более сложный образ. Например, задача формирования внутренних репрезентаций воспринимаемой информации – любой образ, воспринимаемый нашими сенсорами (глазом, ухом,...), при поступлении в мозг формирует картину нейронной активности коры мозга, совершенно не похожую на тот образ, что мы видим или слышим. Другой пример – задача формирования цепочки ассоциаций.

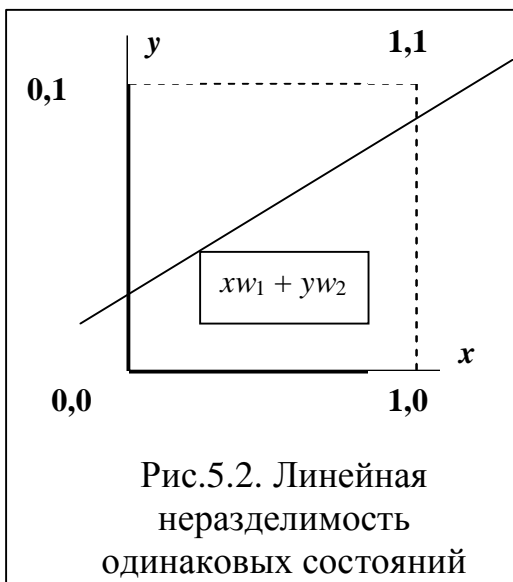
В 1963 г. Розенблатт предложил новый класс моделей мозга, названный им *персептроном* (иногда используется написание *перцептрон*, оба варианта правомочны). Термин персептрон образован от английского perception, или латинского perceptio – восприятие. Первоначально персептрон разрабатывался как нейронная модель зрительного анализатора, в дальнейшем персептрон стал рассматриваться как модель мозга. Разработано множество моделей персептрона, отличающихся друг от друга как топологией связей, так и особенностями отдельных слоев. Мы рассмотрим наиболее общий вариант.

Персептрон, изображенный на рис.5.1, состоит из нескольких слоев нейронов, межнейронные связи внутри слоев отсутствуют, возбуждение распространяется только в одном направлении – от входного слоя к

выходному один раз. Нейроны сенсорного слоя (I_n) никаких вычислений не производят, связи от нейронов сенсорного слоя I_n к нейронам первого вычислительного (А) имеют одинаковые веса, их задача – передать возбуждение от каждого сенсорного нейрона всем нейронам следующего слоя. Собственно обработка информации производится структурой между слоями А и С. Между слоями А и С могут быть еще несколько внутренних (скрытых) слоев (В). Эти слои, включая слой А, называются ассоциативными слоями. Выходной слой (С) иногда называется слоем реакций и используют другую топологию связей V и модель нейрона, отличающуюся от модели нейрона ассоциативного уровня.

В отличие от формального нейрона Мак-Каллока и Питтса, формальный нейрон в некоторых моделях персептрона имеет собственный вес. Время дискретно, т. е. делится на моменты, работа всех элементов синхронизирована. Ассоциативный нейрон срабатывает в момент времени $t + 1$, если сумма сигналов на его входах в момент времени t превышает порог. Пороги срабатывания всех нейронов равны и постоянны. В варианте, когда каждый нейрон имеет свой вес, величина выходного сигнала прямо зависит от веса этого нейрона. Вес нейрона – величина переменная, изменяющаяся при обучении.

Если активационные функции нейронов линейны, то такая многослойная структура может быть сведена к двухслойной сети, состоящей из слоя сенсорных нейронов I_n и одного слоя вычислительных нейронов.



Минский и Пайперт показали возникающую при использовании линейной активационной функции *проблему представимости персептронами как проблему линейной разделимости*. Простейший пример – проблема «исключающего или». Имеется два входных нейрона x и y , принимающих значения 0 и 1 и связанных с вычислительным нейроном связями w_1 и w_2 . Рис.5.2. показывает, что такая сеть не может разделить состояния, когда активирован только один нейрон (неважно какой) от состояний, когда оба

нейрона находятся в одинаковом состоянии. Действительно, выход нейрона $OUT = xw_1 + yw_2$, т.е. в пространстве xu описывает прямую линию, которую невозможно провести так, чтобы по одну сторону остались состояния 0,1 и 1,0, а по другую – 0,0 и 1,1. Если активационные функции линейны, то увеличение количества слоев не решает проблему, т.к. такая сеть сводится к однослойной.

Выход заключается в применении *нелинейных активационных функций*. Линейная разделимость ограничивает однослойные сети задачами классификации, в которых множества точек (соответствующих входным значениям) могут быть разделены геометрически. Двухслойная сеть обеспечивает выпуклую область решений. Трехслойная сеть – вогнутая область решений.

Основная проблема перцептронов – обучение, т.е. формирование структуры связей, обеспечивающей решение задачи. Реализовать в перцептронах правило Хэбба в общем случае непосредственно, как в голографии, нельзя – его можно непосредственно реализовать только в двухслойных перцептронах. Поэтому для обучения многослойных перцептронов применяются методы, относящиеся ко второму классу методов обучения – это *методы коррекции ошибки*. В отличие от правила Хэбба, их уже трудно определить как «биологически мотивированные».

Стохастические методы обучения могут трактоваться как «метод тыка» и обычно включают в себя следующие этапы:

1. Инициализация сети - весам присваиваются произвольные малые значения;
2. сети предъявляется входной вектор, вычисляется выходной, затем
3. производится случайное изменение весов. Если изменение весов приближает выходной вектор к требуемому, то это изменение принимается, если нет, то отвергается.
4. Повторяются шаги 2-3 для всех пар обучающих векторов до тех пор, пока не наступит один из следующих исходов:
 - a. Суммарная ошибка сети по всем парам обучающих векторов не станет меньше заданного порога – сеть считается обученной;
 - b. Сеть перестанет реагировать на любые изменения весов, т.е. войдет в режим насыщения. В этом случае необходимо сеть разобучить, т.е. перейти к шагу 1 и начать все с начала;
 - c. Сеть войдет в циклический режим – состояния выходного слоя станут повторяться с некоторым периодом. В этом случае также необходимо начать процедуру обучения заново.

Поскольку сеть обучается не одной паре векторов, а несколькими парам, то по завершении обучения для одной пары векторов переходят к обучению второй парой, но при этом «портится» результат обучения первой пары. Обучение третьей парой портит результат обучения для первой и второй и т.д. Поэтому по завершении обучения всеми парами векторов, а этот период называется *эпохой*, необходимо вернуться к обучению первой парой, второй, и т.д. Таких эпох может быть много – сотни, тысячи, десятки и сотни тысяч.

Дельта-правило (метод обучения) Розенблатта.

Обучение сети состоит в подстройке весовых коэффициентов связей каждого нейрона. Метод состоит в итерационной подстройке матрицы весов, последовательно уменьшающей ошибку в выходных векторах.

Обучающая выборка – пара векторов (A^α, B^α) , матрица весов $W^\alpha = A^{\alpha T} B^\alpha$.

Весам присваиваются произвольные малые значения случайным образом. При предъявлении необученной сети с произвольной матрицей памяти W входного образа (вектора) A^α формируется выходной вектор $B'^\alpha = A^\alpha W$, а затем вычисляется вектор ошибки $\delta^\alpha = B'^\alpha - B^\alpha$.

Коррекция весов производится по формуле:

$$W_{t+\Delta t} = W_t - h A^{\alpha T} \delta^\alpha,$$

где h – параметр, задающий темп обучения, $h \in [0, 1]$.

Эти шаги повторяются для всей обучающей выборки. Один цикл последовательного предъявления всей выборки называется *эпохой*. Обучение завершается по истечении нескольких эпох в том случае, если:

- а) процесс сойдется, т.е. вектор весов перестает изменяться;
- б) полная, просуммированная по всем векторам абсолютная ошибка станет меньше некоторого значения. Допустимый уровень ошибки определяется задачей.

Обучение НС как задача оптимизации.

Полная ошибка сети в состоянии W определяется как L_2 норма, а именно:

$$E = \sum_{\alpha} \|A^\alpha W - B^\alpha\|^2 = \sum_{\alpha} \sum_i [A^\alpha W - B^\alpha]^2. \quad (5.1)$$

Для решения задачи важны следующие свойства полной ошибки:

а) Ошибка $E = E(W)$ является функцией состояния W , определенной на пространстве состояний. По определению (5.1) она принимает неотрицательные значения.

б) По определению (5.1) функция $E(W)$ принимает нулевое значение в обученном состоянии W^* , которые, тем самым, суть точки минимума функции ошибки (5.1).

Таким образом, задача обучения трансформируется в задачу поиска минимума функции ошибки в пространстве состояний, и, следовательно, для ее решения применимы стандартные методы теории оптимизации.

Для применения метода Розенблатта необходимо знать как текущие выходы нейронов B , так и эталонные значения B . Для многослойной сети эталонные значения известны только для нейронов выходного слоя. Требуемые значения выходов для нейронов скрытых слоев неизвестны.

Обучение методом обратного распространения ошибок.

Для обучения многослойной сети в 1986 г. Руммельхартом и Хинтоном был предложен алгоритм обратного распространения ошибок (*error back propagation*), решающий проблему неизвестных значений состояний скрытых нейронов.

Основная идея состоит в том, что оценка ошибки элементов скрытых слоев есть взвешенная сумма ошибок последующих слоев. При обучении информация распространяется от входного слоя к выходному, а оценки ошибок – в обратном направлении, что и отражено в названии метода.

Обозначения:

В соответствии с рис.5.1, примем следующие обозначения:

- матрица весовых коэффициентов от входов к внутреннему слою W , а матрица весов, соединяющих скрытый и выходной слой – V .
- нейроны входного слоя нумеруются индексом i ;
- элементы скрытого слоя – индексом j ;
- выходы нейронов, соответственно, индексом k ;
- обучающая выборка (A^α, C^α) , $\alpha = 1..p$.

Состояния нейронов (выходы) обозначим маленькой буквой слоя с соответствующим индексом, а суммарные взвешенные входы нейронов – x с соответствующим нижним индексом.

Общая структура алгоритма обучения аналогична вышеописанному алгоритму Розенблатта:

1. начальные значения всех весов выбираются случайным образом;
2. предъявляется образ A^α , на выходе формируется $C^{\alpha'} \neq C^\alpha$.

Для описания состояния нейронов примем следующие обозначения:

слой В (скрытый):	входы нейронов:	$x_j = \sum_i W_{ij} a_i^\alpha,$
	выходы нейронов:	$b_j = f(x_j);$
слой С (выходной):	входы нейронов:	$x_k = \sum_j V_{jk} b_j^\alpha,$
	выходы нейронов:	$c_k = f(x_k).$

Подлежащий минимизации функционал ошибки сети определим следующим образом:

$$E = \frac{1}{2} \sum_k (c_k - c_k^\alpha)^2.$$

Градиентный метод оптимизации заключается в коррекции весов связей по формуле:

$$v_{jk}(t+1) = v_{jk}(t) - h \frac{\partial E_k}{\partial v_{jk}}, \tag{5.2}$$

где параметр h имеет смысл темпа обучения. Поскольку функция ошибки не содержит в явном виде зависимости от V , то представим ошибку в виде:

$$\frac{\partial E_k}{\partial v_{jk}} = \frac{\partial E_k}{\partial c_k} \frac{\partial c_k}{\partial x_k} \frac{\partial x_k}{\partial v_{jk}},$$

где:

$$\frac{\partial E_k}{\partial c_k} = (c_k - c_k^\alpha) = \delta_k,$$

$$\frac{\partial c_k}{\partial x_k} = c_k(1 - c_k),$$

поскольку производная сигмоидальной функции $f(x)$ выражается только через само значение функции, $f'(x) = f(1-f)$. Тогда:

$$\frac{\partial x_k}{\partial v_{jk}} = b_j.$$

Отсюда получим окончательное выражение для производной ошибки:

$$\frac{\partial E_k}{\partial v_{jk}} = \delta_k c_k b_j (1 - c_k),$$

т.е. формула подстройка весов выходного слоя имеет вид:

$$v_{jk}(t+1) = v_{jk}(t) - h \delta_k c_k b_j (1 - c_k).$$

3. Следующий шаг – подстройка весов скрытого слоя (В) аналогичным методом, т.е.:

$$w_{ij}(t+1) = w_{ij}(t) - h \frac{\partial E_j}{\partial w_{ij}}.$$

Формула для расчета ошибки:

$$\frac{\partial E_j}{\partial w_{ij}} = \frac{\partial E}{\partial b_j} \frac{\partial b_j}{\partial x_j} \frac{\partial x_j}{\partial w_{ij}},$$

где первая производная может быть представлена в виде:

$$\frac{\partial E_j}{\partial b_j} = \delta_j = \sum_k \frac{\partial E_k}{\partial x_k} \frac{\partial x_k}{\partial b_j} = \sum_k \frac{\partial E_k}{\partial c_k} \frac{\partial c_k}{\partial x_k} \frac{\partial x_k}{\partial b_j} = \sum_k \delta_k c_k (1 - c_k) v_{jk}.$$

Отсюда получим окончательное выражение для правила подстройки весов скрытого слоя:

$$\frac{\partial E_j}{\partial w_{ij}} = \frac{\partial E_j}{\partial b_j} \frac{\partial b_j}{\partial x_j} \frac{\partial x_j}{\partial w_{ij}} = \left(\sum_k \delta_k c_k (1 - c_k) v_{jk} \right) (b_j (1 - b_j) a_j^\alpha).$$

Принцип обратного распространения ошибки заключается в том, что частные производные при вычислении δ_j берутся только по переменным последующего слоя, причем оценка ошибки для каждого слоя суть взвешенная сумма ошибок последующего слоя. Если в сети несколько скрытых слоев, то такая процедура применяется последовательно для каждого из них. Эти шаги повторяются для всей обучающей выборки.

Суть метода обратного распространения ошибки выражается формулой:

$$\frac{\partial E_j}{\partial b_j} = \delta_j = \sum_k \frac{\partial E_k}{\partial x_k} \frac{\partial x_k}{\partial b_j} = \sum_k \frac{\partial E_k}{\partial c_k} \frac{\partial c_k}{\partial x_k} \frac{\partial x_k}{\partial b_j} = \sum_k \delta_k c_k (1 - c_k) v_{jk}.$$

Параметр h в выражении (5.2) (темп обучения) для обеспечения сходимости метода должен быть небольшим. В противном случае сеть может «перескочить» через требуемое состояние.

Особенности метода следующие:

1. В силу того, что ошибка может иметь достаточно сложную, не гладкую поверхность, локальное направление градиента отнюдь не всегда совпадает с направлением к минимуму. Сеть может долго «блуждать» по оврагам и долинам, которыми изрезана поверхность ошибки, прежде чем

«выйдет» к глобальному минимуму.

Поэтому сходимость метода, как и большинства градиентных методов, медленная.

2. Поскольку подстройка весов выполняется независимо для каждой пары образов обучающей выборки, уменьшение ошибки для какой-либо пары может «портить» результат обучения для предыдущих образов. В принципе, гарантий сходимости метода нет.

Разработано множество модификаций метода, нацеленных на решение тех или иных частных задач. Но общая идея остается старой. Единственно, на чем необходимо остановить внимание – на *проблеме*



выхода из локальных минимумов.

Поверхность ошибки может быть не гладкой, но обладать локальными минимумами, при попадании в который процесс обучения остановится в некотором, отличном от требуемого, состоянии (Рис.5.3). Для решения проблемы локальных минимумов предложен ряд методов, основная идея которых заключается в том, что с некоторой вероятностью сеть может не

только спускаться по поверхности ошибки, но и подниматься для того, чтобы выйти из локального минимума. По мере обучения вероятность подъема уменьшается. Поэтому эти методы известны также как стохастические методы.

К числу таких методов относятся так называемые «термодинамические» методы или методы «отжига» – методы, основанные на использовании аналогий (на уровне математических моделей) с процессами, протекающими в термодинамических системах и вводящие понятие искусственной температуры сети, которая по мере обучения понижается. Если представить состояние сети в виде шарика, скатывающегося по поверхности ошибки, то температура, т.е. внутренняя энергия сети, позволяет шарика «подпрыгивать» и, тем самым, «выпрыгивать» из локальных минимумов.

Например, метод обучения «машина Больцмана» предполагает, что в случае, если изменение веса приводит к увеличению ошибки, то в отличие от классических методов, запрещающих такие изменения веса, в данном методе такие изменения допускаются с вероятностью

$$P(E) = \exp\left(-\frac{\Delta E}{kT}\right) > r,$$

где k – постоянная Больцмана, T – искусственная температура сети, ΔE – изменение ошибки, $r \in [0,1]$ – порог. Скорость уменьшения искусственной температуры для обеспечения сходимости процесса у «машины Больцмана» определяется выражением:

$$T(t) = \frac{T_0}{\lg(1+t)},$$

где t -искусственное время.

Алгоритм обучения «машина Коши» отличается от «машины Больцмана» тем, что вероятность изменения веса, увеличивающего ошибку, определяется выражением для распределения Коши:

$$P(E) = \frac{T(t)}{T(t)^2 + E^2}.$$

В этом случае время обучения сокращается согласно выражению:

$$T(t) = \frac{T_0}{1+t}.$$

Эффект ускорения обучения достигается за счет того, что распределение Коши обеспечивает большую вероятность больших шагов по сравнению с распределением Больцмана (Рис.5.4).

Известны также и другие модели, отличающиеся используемым распределением, но общая идея у всех одна и та же.

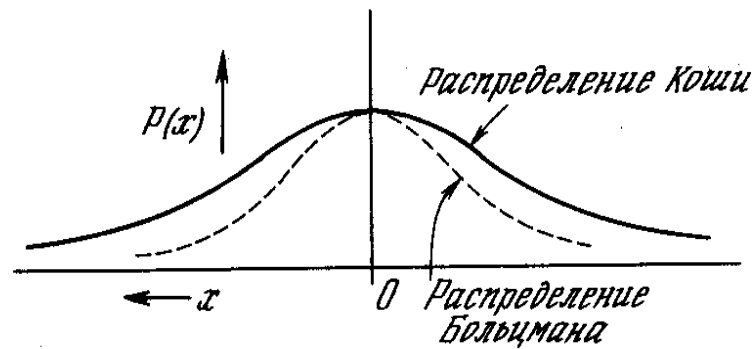


Рис.5.4. Распределение Коши и распределение Больцмана [2]

В развитие идеи обучения методом *коррекции ошибки* с использованием *градиентной оптимизации* разработано множество методов. В настоящее время наибольший интерес представляют комбинированные методы, включающие два этапа:

- на первом этапе производится глобальный поиск области минимума ошибки – здесь хорошо зарекомендовали себя методы, использующие генетические алгоритмы, которые позволяют быстро локализовать область существования экстремума, но не позволяют найти точное решение;
- на втором этапе производится поиск уточненного решения методом градиентного спуска в определенной на первом этапе локальной области.

Оптические реализации.

Основная проблема реализации персептронов «в железе» заключается в обеспечении полного набора связей – каждый нейрон должен быть связан со всеми нейронами последующего слоя. Это требование выполняется в оптике за счет использования явления дифракции – каждый нейрон как точечный источник формирует расходящуюся сферическую волну, освещающую все нейроны следующего слоя. Распространение оптического излучения в свободном пространстве обеспечивает независимость связей (невзаимодействие волн от каждого нейрона – точечного источника). В то же время, каждая связь должна иметь свой вес. Веса связей могут быть реализованы в виде функции пропускания или дифракционной эффективности, т.е. необходим материальный носитель.

Для реализации нейронных слоев и матриц весов связей наиболее часто в настоящее время используются пространственно-временные модуляторы света на основе структур фотополупроводник – жидкий кристалл. Эти структуры могут работать как под электрическим управлением (ЖК дисплей), так и под оптическим. Оптически управляемые ЖК модуляторы света используются для записи голограмм, что позволяет реализовать важнейший НС принцип “model free” – сеть обучается (записывается голограмма) при предъявлении изображения, при этом мы не знаем что это за изображение, каковы его свойства, не нуждаемся в его формальном описании. Но принцип “model free”, как и всякий базовый принцип, в

реальности ограничен тем, что мы должны обеспечить правильный подбор условий экспозиции!

Ограничения оптических технологий.

Голография позволяет реализовать правило обучения Хебба, но с ограничениями:

Обучение только одного слоя;

Структура связей. Если число нейронов во входном слое n (по одной координате), то в пространстве Фурье:

$$x = \lambda f v_{\max}, \quad d = \frac{\lambda f}{D}.$$

Отсюда получаем число элементов:

$$\frac{2x}{d} = \frac{2\lambda f v_{\max} D}{\lambda f} = \frac{2D}{2\Delta x} = n,$$

т.е. полный набор связей возможен только по схеме $Im \leftrightarrow \delta$, где δ – нейрон, активированный изображением дифракционно-ограниченного точечного опорного источника (глобального максимума АКФ).

Если образы представлены как изображения, т.е. область их определения имеет размерность $2D$, то размерность области определения матрицы связей будет $4D$, что физически нереализуемо. Таким образом, реализуемы схемы обработки $1-D$ векторов – вектор-матричные умножители. При использовании объемных (трехмерных) регистрирующих сред реализуемы матрицы связей вектора-строки и $2-D$ матрицы. Если число связей на нейрон ограничено, то реализуемы матрицы связей двух $2-D$ матриц посредством пространственной локализации каждой субматрицы связей.

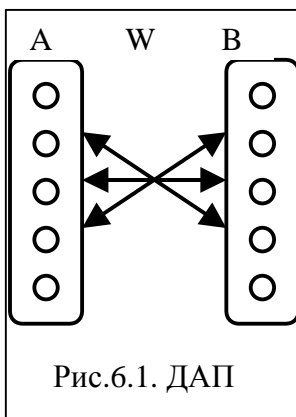
Литература к Лекции 5

1. *Кольер Р., Беркхарт Л., Лин Л.*, "Оптическая голография", М.Мир, 1973, 686с.
2. *Уоссермен Ф.*, Нейрокомпьютерная техника - М., Мир, 1992.
3. *Комарцова Л.Г., Воеводин Ю.Ю.*, Исследование комбинированных алгоритмов обучения нейронной сети для решения задач классификации // Сб. научных трудов III Межд. Научно-практического семинара «Интегрированные модели и мягкие вычисления в искусственном интеллекте». – М.:Физматлит. –2005. – С.306–308.

Лекция 6. Искусственные нейронные сети как динамические системы

В персептронах информация распространяется от входа к выходу один раз, т.е. персептрон работает по принципу «вопрос – ответ». В то же время, процесс мышления зачастую представляет собой долгие размышления. Эту сторону мышления моделируют НС, относящиеся к категории динамических систем. В этих моделях состояние сети, активированной поступившим на ее вход вектором, изменяется в результате взаимодействия нейронов.

Двунаправленная ассоциативная память (ДАП)



ДАП была предложена Б.Коско и состоит из двух связанных между собой нейронных слоев A и B, соединенных матрицей связей W (рис.6.1.). Внутри слоев межнейронные связи отсутствуют. Если состояние слоя описывается вектором-строкой A и B, соответственно, то матрица связей определяется как их произведение, т.е.

$$W = A^T B.$$

Простейший вариант ДАП – сеть с бинарными нейронами, выходы которых определяются по правилу

$$a_i = \begin{cases} 1 & \text{if } BW_i^T > 0 \\ 0 & \text{if } BW_i^T < 0 \end{cases} \quad (6.1)$$

$$b_i = \begin{cases} 1 & \text{if } AW_i > 0 \\ 0 & \text{if } AW_i < 0 \end{cases} \quad (6.1)$$

Если сеть обучена паре векторов A_1, B_1 , то при поступлении на слой A вектора A_1 , на слой B через матрицу связей поступает возбуждение $A_1 W = A_1 A_1^T B_1 = B_1$, т.е. в слое B восстанавливается эталонный вектор B_1 .

Если в слой A поступает вектор A' , отличный от эталонного, то в слое B восстанавливается $A' W = A' A_1^T B_1 = B'$, который проходит через матрицу связей в обратном направлении и в слое A формирует вектор:

$$B' W^T = B' B_1^T A_1 = A''_1,$$

который снова проходит через матрицу связей и т.д., т.е. развивается итерационный процесс.

Для определения направления этого процесса и возможность его схождения к некоторому устойчивому состоянию, введем функцию, которая при любом изменении состояния сети не будет возрастать и будет иметь единственный минимум в точке, к которой должен сойтись процесс – функцию Ляпунова или энергию сети. В общем случае нахождение такой

функции является сложной задачей, не имеющей наперед заданного способа решения. Для модели ДАП энергия сети может быть определена как скалярное произведение ассоциируемых векторов, т.е. в виде:

$$E(A, B) = -\frac{1}{2}AWB^T - \frac{1}{2}BW^T A^T = -AWB^T = -\sum_i \sum_j a_i b_j w_{ij}, \quad (6.2)$$

а с учетом порогов нейронов T , S и их входов I , J выражение для энергии примет вид:

$$E(A, B) = -AWB^T - IA^T + TA^T - JB^T + SB^T$$

Изменение энергии сети при произвольном изменении состояния любого нейрона Δa_i :

$$\Delta E(A, B) = -\Delta A W B^T = -\sum_i \Delta a_i B W^T. \quad (6.3)$$

Можно видеть, что в соответствии с принятыми условиями (6.1), при любом изменении состояния нейрона энергия сети не возрастает. Рассмотрим все возможные варианты суммарного входа нейрона a_i слоя A :

1. Пусть на вход нейрона поступает $WB^T > 0$. Тогда, в соответствии с (6.1), нейрон a_i должен перейти в состояние 1, т.е. если он был в состоянии 0, то $\Delta a_i = 1$, т.е. $\Delta a_i > 0$ и, в соответствии с (6.2), энергия сети уменьшается. Если нейрон был в состоянии 1, то $\Delta a_i = 0$, и в соответствии с (6.2) энергия сети осталась прежней.
2. Пусть $WB^T < 0$. Тогда, в соответствии с (6.1), нейрон a_i должен перейти в состояние 0, т.е. если он был в состоянии 1, то $\Delta a_i = -1$, т.е. $\Delta a_i < 0$, в соответствии с (6.2) энергия сети уменьшается. Если нейрон был в состоянии 0, то $\Delta a_i = 0$, и в соответствии с (6.4) энергия сети осталась прежней.
3. Если $WB^T = 0$, то нейрон a_i не изменяет свое состояние, энергия сети осталась прежней.

Таким образом, при любом возможном изменении состояния нейрона энергия сети не возрастает. В результате, в пространстве состояний сети функция энергии образует поверхность с минимумом, соответствующим паре эталонных векторов – аттрактором. Следовательно, будучи после поступления входного вектора предоставленной самой себе, сеть сойдется (релаксирует) к состоянию с минимальной энергией в котором и будет резонировать. Если энергия определена как скалярное произведение векторов, а матрица весов как внешнее произведение векторов, то минимум энергии соответствует паре эталонных векторов $A \leftrightarrow B$. Таким образом, ДАП Коско реализует модель *ассоциативной памяти*. Если сети предъявляется вектор A' , а вектор, к которому сходится сеть в результате итерационного процесса, выводится из слоя A , то сеть реализует модель *авто-ассоциативной памяти*. Если результат выводится из слоя B , то реализуется *гетеро-ассоциативная память*.

Сети Хопфилда – сети полной связности

В сетях Хопфилда отсутствует «слоистая» структура и на вход каждого нейрона поступают выходы всех нейронов, в том числе, в некоторых моделях и собственный выход. Сеть Хопфилда реализует модель *автоассоциативной памяти*.

Для демонстрации принципа работы сети Хопфилда рассмотрим простейшую модель с бинарными нейронами, как это сделано выше для ДАП Коско. Определим матрицу связей аналогично тому, как определяли ее и раньше, т.е. как внешнее произведение векторов:

$$W = A^T A. \quad (6.4)$$

Пусть j -ый нейрон с порогом T_j , его состояние NET_j определяется как сумма входов:

$$NET_j = \sum_{i \neq j} w_{ij} OUT_i + IN_j. \quad (6.5)$$

Примем правило формирования выхода нейрона, аналогичное (6.1), уже использованному при рассмотрении модели ДАП, а именно:

$$OUT_j = 1, \text{ если } \sum_{i \neq j} w_{ij} OUT_i + IN_j > T_j,$$

$$OUT_j = 0, \text{ если } \sum_{i \neq j} w_{ij} OUT_i + IN_j < T_j,$$

$$OUT \text{ не изменяется, если } \sum_{i \neq j} w_{ij} OUT_i + IN_j = T_j.$$

Энергию сети Хопфилда (функцию Ляпунова) также определим аналогично ранее использованному, т.е. как скалярное произведение:

$$E = -\frac{1}{2} \sum_i \sum_j w_{ij} OUT_i OUT_j - \sum_j IN_j OUT_j + \sum_j T_j OUT_j, \quad (6.6)$$

где w_{ij} – вес от выхода нейрона i к входу нейрона j ; OUT_j – выход нейрона j ; IN_j – внешний вход нейрона j ; T_j – порог нейрона j .

Изменение энергии E , вызванное изменением состояния j -нейрона:

$$\delta E = - \left[\sum_{i \neq j} (w_{ij} OUT_i) + IN_j - T_j \right] \delta OUT_j = - [NET_j - T_j] \delta OUT_j, \quad (6.7)$$

где δOUT_j – изменение выхода j -го нейрона. Если имеет место:

$$\sum_{i \neq j} w_{ij} OUT_i + IN_j > T_j \Rightarrow \delta E > 0,$$

то и OUT тоже должен измениться в большую сторону или остаться без изменения, следовательно, энергия уменьшается. Если имеет место:

$$\sum_{i \neq j} w_{ij} OUT_i + IN_j < T_j \Rightarrow \delta E < 0,$$

то и ОУТ тоже должен измениться в меньшую сторону или остаться без изменения, следовательно, энергия уменьшается.

Таким образом, как и в случае ДАП, любое изменение состояния нейрона ведет к уменьшению энергии сети – сеть «скатывается» в точку минимума энергии. Задача обучения заключается в том, чтобы точка минимума энергии совпадала с эталонным состоянием. Один из вариантов обучения, обеспечивающий выполнение этого условия – использование модели внешнего произведения векторов. *Матрица связей определяется как внешнее произведение эталонных векторов:*

$$W = \sum_i A_i^T A_i,$$

а энергия сети – как внутреннее (скалярное) произведение. Рассмотренные выше примеры (ДАП и сеть Хопфилда) полностью укладываются в эти рамки.

Также нетрудно убедиться в том, что в эти рамки полностью укладывается и рассмотренная выше модель голографической автоассоциативной памяти, реализуемая в схеме «Голографический коррелятор в линейном резонаторе».

Устойчивость сети. В приведенных примерах сетей с бинарными нейронами динамические системы устойчивы – матрицы связей, сформированные по правилу внешнего произведения, обеспечивают наличие в пространстве состояний для каждой запомненной пары (для ДАП) или вектора (для НС Хопфилда) единственного минимума функции энергии – аттрактора, к которому сходится сеть. Если используется другая модель формирования матрицы связей, то наличие единственного аттрактора не гарантировано и сеть может перейти в режим хаотической динамики или замкнутый цикл. Эти режимы также представляют интерес с точки зрения задачи создания искусственного интеллекта и будут более подробно обсуждены в следующих лекциях данного курса. Сейчас лишь укажем, что в общем случае для обеспечения устойчивости системы достаточно, чтобы матрица связей была симметрична и имела нули на главной диагонали, т.е.:

$$\begin{aligned} w_{ij} &= w_{ji} \Big|_{i \neq j} \\ w_{ij} &= 0 \Big|_{i=j} \end{aligned}.$$

Задача оптимизации. НС Хопфилда может использоваться не только для реализации модели авто-ассоциативной памяти, но и для решения задач комбинаторной оптимизации, отличающихся от задач градиентной оптимизации. Задача комбинаторной оптимизации заключается в нахождении наилучшего решения из числа конечного, но, как правило, очень большого возможных. Классический пример задачи комбинаторной оптимизации – «задача коммивояжера», заключающаяся в том, что

необходимо объехать N городов так, чтобы каждый город был посещен только один раз, а длина маршрута была минимальна. Единственный метод, гарантирующий решение таких задач – метод перебора. Нетрудно видеть, что если размерность задачи N (т.е. число городов), то число возможных решений $N!$, т.е. задача решается за экспоненциальное время, поскольку формула Стирлинга для достаточно больших N дает $N! \approx \left(\frac{N}{e}\right)^N$.

Приемлемыми считаются методы, дающие решение за полиномиальное время, т.е. время, растущее с ростом размерности задачи как N . Если возможно нахождение экстремума целевой функции за полиномиальное время, то задача относится к классу P. NP класс образуют задачи, для которых за полиномиальное время можно оценить значение целевой функции для заданной конфигурации, но нельзя перебрать все конфигурации. Наиболее сложные задачи из класса NP называют NP-полными, поскольку если бы удалось найти решение одной из них за полиномиальное время (или доказать его существование), то такое решение существовало бы и для остальных NP-полных задач. Иными словами, все NP-полные задачи сводятся одна к другой за полиномиальное время и методы решения, разработанные для одной из них, можно применить ко всем остальным. Задача коммивояжера – классический пример NP-полной задачи. Хопфилд и Танк показали, что НС Хопфилда может быть использована для решения задач комбинаторной оптимизации - найденное решение не является наилучшим, но удовлетворительным.

Например, для задачи посещения всего 10 городов существует 181440 возможных маршрутов. Попытка решения задачи на основе того подхода, что будет рассмотрен ниже, т.е. с использованием сети Хопфилда, показала, что 16 раз из 20 сеть сходилась к допустимым решениям (маршрутам), из которых порядка 50% оказались кратчайшими.

Идея подхода заключается в конструировании функции энергии сети таким образом, чтобы в точке минимума этой функции выполнялось условие задачи. Пусть имеется N городов, решение представим в виде таблицы 6.1.

Таблица 6.1.

Город	Порядок посещения			
N_1	1	0	0	0
N_2	0	0	0	1
N_3	0	1	0	0
N_n	0	0	1	0

В таблице 6.1 каждый город представлен строкой, номер столбца означает порядок его посещения. Очевидно, что каждый столбец и каждая строка должны содержать только одну отметку о посещении – единицу. Остальные столбцы в строке и строки в столбце должны содержать нули – каждый город посещается только один раз.

Рассмотрим сеть Хопфилда, из $N \times N$ нейронов, которой каждый нейрон имеет два индекса, соответствующих номеру города и номеру его посещения в маршруте. Например, нотация $OUT_{xj} = 1$ означает, что город x был j -ым по порядку городом маршрута. В соответствии с вышеприведенной таблицей в состоянии решения должны быть активны только N нейронов.

Введем функцию энергии сети, обеспечивающую выполнение требования на однократность посещения и охват всех городов:

$$E = \frac{A}{2} \sum_x \sum_i \sum_{j \neq i} OUT_{xi} OUT_{xj} + \frac{B}{2} \sum_i \sum_x \sum_{y \neq x} OUT_{xi} OUT_{yi} + \frac{C}{2} \left[\left(\sum_x \sum_i OUT_{xi} \right) - N \right]^2. \quad (6.8)$$

где A , B и C – коэффициенты. Нетрудно видеть, что в выражении (6.8):

1. Первое слагаемое обеспечивает выполнение требования на посещение каждого города не более одного раза – только в случае, если строка содержит не более одной единицы слагаемое равно нулю
2. Второе слагаемое обеспечивает выполнение требования на посещение только одного города в каждый момент времени – только в случае, если столбец содержит не более одной единицы слагаемое равно нулю
3. Третье слагаемое обеспечивает выполнение требования на посещение всех городов – оно равно нулю только если матрица содержит ровно N единиц.

Теперь необходимо сформулировать требование на предпочтение коротких маршрутов, а именно:

$$E = \frac{D}{2} \sum_x \sum_{y \neq x} \sum_i d_{xy} OUT_{xi} (OUT_{y,i+1} + OUT_{y,i-1}). \quad (6.9)$$

Этот член представляет собой длину любого допустимого маршрута. Здесь индексы определены по модулю N , т. е. $OUT_{N+j} = OUT_j$, D – коэффициент.

При достаточно больших значениях A , B и C низкоэнергетические состояния будут представлять допустимые маршруты, а большие значения D гарантируют, что будет найден короткий маршрут.

Теперь осталось определить веса, т. е. установить соответствие между членами в функции энергии и членами общей формулы для энергии сети Хопфилда. Получаем:

$$\begin{aligned} w_{xi,yi} &= -A\delta_{xy}(1 - \delta_{ij}) && \text{(не допускает более одной единицы в строке)} \\ &\quad -B\delta_{ij}(1 - \delta_{xy}) && \text{(не допускает более одной единицы в столбце)} \\ &\quad -C && \text{(глобальное ограничение)} \\ &\quad -Dd_{xy}(\delta_{j,i+1} + \delta_{j,i-1}) && \text{(член, отвечающий за длину цикла),} \end{aligned}$$

где $\delta_{ij} = 1$, если $i = j$, в противном случае $\delta_{ij} = 0$. Кроме того, каждый нейрон имеет смещающий вес x_i , соединенный с $+1$ и равный Cn .

Тема 6.2. Самоорганизация и самообучение в НС. Сеть Кохонена

Рассмотренные выше модели НС реализуют идеологию «обучения с учителем», в рамках которой сеть обучается на примерах «вход – желаемый отклик». В то же время, один из атрибутов интеллекта – способность *самостоятельно* формировать *индивидуальную* систему классификации входной информации. Задачу самостоятельного формирования сетью индивидуальной системы классификации решает *сеть Кохонена*. Сеть состоит из двух слоев – сенсорного, не выполняющего никаких вычислений, и собственно *слоя Кохонена*.

Слой Кохонена классифицирует входные векторы в группы подобных. В простейшей форме слой Кохонена реализует модель «победитель забирает все». Эта модель известна также под своим английским названием *WTA – Winner Takes All*.

Можно выделить несколько этапов в работе слоя (сети) Кохонена:

1. Инициализация сети – весам присваиваются случайным образом выбранные малые значения.
2. Самообучение сети (классификация) – при поступлении входного вектора сеть самостоятельно подстраивает веса связей.
3. Дообучение сети – в процессе функционирования при поступлении новых векторов сеть изменяет веса связей так, чтобы они наилучшим образом соответствовали распределению входных векторов.

Состояние j -го нейрона Кохонена определяется, как и в ранее рассмотренных моделях, как сумма его входов:

$$NET_j = \sum_i a_i w_{ij},$$

или в векторной форме $N = AW$, где N – вектор состояний NET слоя Кохонена.

Нейрон с максимальным значением NET назначается «победителем». Его выход равен единице, у остальных он становится равен нулю, т.е. только нейрон-победитель возбужден и формирует выходной сигнал – в этом и заключается идея *WTA*. Для реализации этой идеи применяется механизм *латерального торможения* – внутри слоя Кохонена вводятся два вида связей между нейронами слоя (Рис.6.2.1.):

- каждый нейрон подает на входы других нейронов слоя Кохонена сигналы отрицательной полярности, т.е. тормозящие, пропорциональные уровню своего возбуждения, тем самым снижая уровень возбуждения других нейронов;
- одновременно каждый нейрон подает на свой собственный вход положительный сигнал, пропорциональный уровню своего возбуждения, поддерживая, тем самым, свой уровень возбуждения.

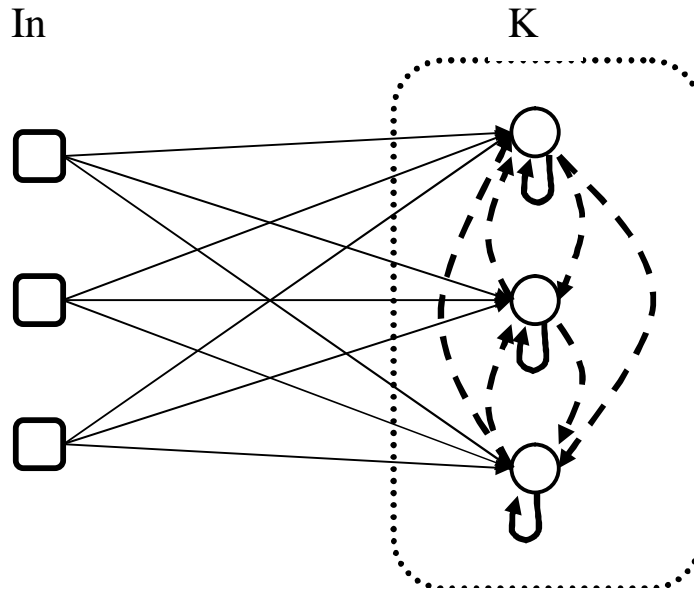


Рис.6.2.1. Сеть Кохонена. In – слой сенсорных нейронов, K – слой Кохонена. Сплошными линиями показаны возбуждающие связи, пунктиром – тормозящие

В результате развивается процесс взаимодействия нейронов в слое Кохонена, приводящий к выявлению нейрона с максимальным уровнем возбуждения – нейрона победителя, который затормаживает все остальные нейроны.

Инициализация сети. Весам связей от слоя сенсорных нейронов к нейронам слоя Кохонена присваиваются случайным образом выбранные малые значения.

Обучение слоя Кохонена. Заранее неизвестно какой именно нейрон Кохонена будет активироваться для заданного входного вектора. Единственное условие - должно быть обеспечено разделение векторов, т.е. несхожие входные векторы должны активировать разные нейроны-победители.

Первый (предварительный) этап - нормализация входных векторов:

$$a'_i = \frac{a_i}{\sqrt{a_1^2 + a_2^2 + \dots + a_n^2}}.$$

Нормализация перед началом обучения приближает веса к их окончательным значениям, сокращая, тем самым, время обучения.

Далее:

- a. вводится метрика, например, скалярное произведение;
- b. вычисляются расстояния входного вектора с векторами весов;
- c. нейрон с максимальным значением скалярного произведения объявляется «победителем» и его веса подстраиваются по формуле:

$$\omega(t+1) = \omega(t) + h(a - \omega(t)),$$

где $\omega(t+1)$ – новое значение веса, соединяющего входную компоненту a с выигравшим нейроном; $\omega(t)$ – предыдущее значение этого веса; h –

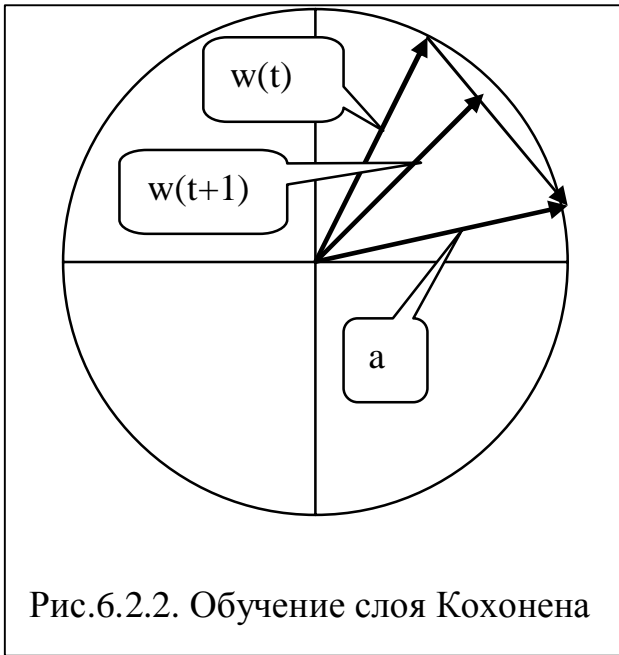


Рис.6.2.2. Обучение слоя Кохонена

коэффициент скорости обучения, который может варьироваться в процессе обучения.

Каждый вес, связанный с нейроном – победителем, изменяется пропорционально разности между его величиной и величиной входа, с которым он ассоциирован. Направление изменения минимизирует разность между весом и его входом.

Если с каждым нейроном Кохонена ассоциируется только один входной вектор, то для обучения достаточно одного вычисления на вес ($h = 1$). Если

имеется набор обучающих векторов которые должны быть отнесены к одному классу, то сеть должна активировать один и тот же нейрон для всех этих векторов, т.е. веса этого нейрона должны получаться усреднением группы входных векторов. Постепенное уменьшение h уменьшает влияние каждого обучающего шага, в результате окончательное значение будет средней величиной от обучающих векторов. Таким образом, веса, ассоциированные с нейроном, сгруппируются вблизи «центра» входных векторов, для которых данный нейрон является «победителем».

Проблема – как распределить начальные значения весовых векторов в соответствии с плотностью распределения входных векторов, которые должны быть разделены? Одно из приближенных решений – метод выпуклой комбинации (convex combination method). Все веса приравниваются согласно формуле:

$$w_i = \frac{1}{\sqrt{n}},$$

где n – число входов, т.е. число компонент каждого весового вектора. Благодаря этому все весовые вектора совпадают и имеют единичную длину. Каждой же компоненте входа A придается значение:

$$a_i = \alpha a_i + \frac{1-\alpha}{\sqrt{n}},$$

где n – число входов. В начале обучения параметр α очень мал, в результате длина всех обучающих векторов близка к $\frac{1}{\sqrt{n}}$, т.е. вектора почти совпадают с векторами весов. В процессе обучения сети α постепенно увеличивается до единицы. Это ведет к разделению входных векторов с приписыванием им их истинных значений. Весовые векторы отслеживают один или небольшую группу входных векторов и в конце

обучения дают требуемую картину выходов. Метод выпуклой комбинации существенно замедляет процесс обучения, поскольку подстройка весов производится к цели, изменяющейся во время обучения.

Другие методы, используемые для распределения начальных значений весов в соответствии с возможным распределением векторов:

Добавление шума к входным векторам.

На начальной стадии подстраиваются все веса, а затем коррекция весов производится только для нейронов, ближайших к победителю.

Введение в модель «Чувства справедливости». Если какой-либо нейрон становится победителем слишком часто (чаще, чем $1/k$, где k – число нейронов Кохонена), то он временно увеличивает свой порог, давая тем самым возможность обучаться и другим нейронам.

Если для каждого входного вектора активируется лишь один нейрон-победитель, то такой режим называется методом аккредитации (назначения). Недостаток метода аккредитации в ограниченности его точности, так как выход слоя суть функция только одного нейрона. Модификация идеи WTA – возбуждение не одного, а группы нейронов с максимальным уровнем активации – метод интерполяции. Поскольку размерность выходного вектора увеличивается, то возрастают и интеллектуальные способности такой модели.

Оптические реализации слоя Кохонена.

Нетрудно видеть, что ключевым моментом в построении слоя Кохонена является реализация идеологии WTA. Один из возможных вариантов реализации этой идеи – использование для реализации слоя Кохонена оптических сред с нелинейными передаточными характеристиками, подобранными таким образом, чтобы в результате того или иного физического процесса, протекающего в среде, по его завершении оставался возбужденным единственный нейрон – победитель.

Литература к Лекции 6.2.

1. Уоссермен Ф., Нейрокомпьютерная техника – М.: Мир, 1992.

Лекция 7. Концепция адаптивного резонанса (Концепция ART)

Проблема стабильности – пластичности памяти.

Для персептрона обучение производится сразу для всего набора запоминаемых образов и попытка дообучения обученной сети новому образу может разрушить всю структуру памяти и может потребоваться полностью повторить всю процедуру обучения. В то же время любой живой организм в процессе своей жизнедеятельности постоянно пополняет свою память новыми образами, не разрушая ее структуру, но сохраняя в памяти уже запомненные образы, частично их модифицируя.

Суть проблемы (или дилеммы) *стабильности-пластичности памяти* в том, что НС при восприятии новой информации должна не просто добавить ее в память, но соотнести новую информацию с уже запомненной и, если есть какое-то сходство между новой и старой информацией, то скорректировать информацию запомненную, т.е. «принять во внимание» новую информацию. Один из подходов к решению этой проблемы был предложен Гроссбергом и Карпенгером в виде концепции ART – *Теории Адаптивного Резонанса*.

В настоящее время идея ART конкретизирована в целом ряде моделей, объединенных тем общим, что позволяет определить ART не просто как одну из моделей нейронных сетей, но как нечто большее – как микропарадигму в НС парадигме. Для понимания общей идеи рассмотрим простейшую модель – ART1, работающую с бинарными векторами.

В ART введено нескольких существенных элементов, а именно:

Наличие внутреннего «*детектора новизны*» – если критерий новизны удовлетворяется, то сеть классифицирует предъявленный образ как новый, если критерий не удовлетворяется, то возникает резонанс на моде, соответствующей одному из ранее запомненных образов, при этом сеть изменяет запомненный ранее образ, т.е. резонанс адаптивный.

Вводится «*шаблон критических черт*» (critical feature pattern) информации. В идеале, выделение критической части информации должно происходить автоматически в процессе работы и обучения сети, на основе ее индивидуального опыта.

«*Ориентирующая подсистема*» сама принимает решение о прекращении поиска в памяти и переходе к обучению новому вектору. Ориентирующая подсистема обеспечивает самоадаптацию алгоритма поиска образов в памяти, поскольку НС работает в постоянно изменяющихся условиях.

Правило 2/3 (два из трех) – каждый элемент (нейрон или вентиль) имеет три входа; выходной сигнал формируется при наличии сигналов на двух входах.

Встречное ожидание – сеть находит в памяти ранее запомненный вектор – тот, который «по ее мнению» наиболее соответствует входному,

извлекает его из памяти, и сравнивает с входным, чтобы проверить по заданному критерию сходства действительно ли он соответствует входному, или сеть «ошиблась» и надо поискать в памяти другой вектор.

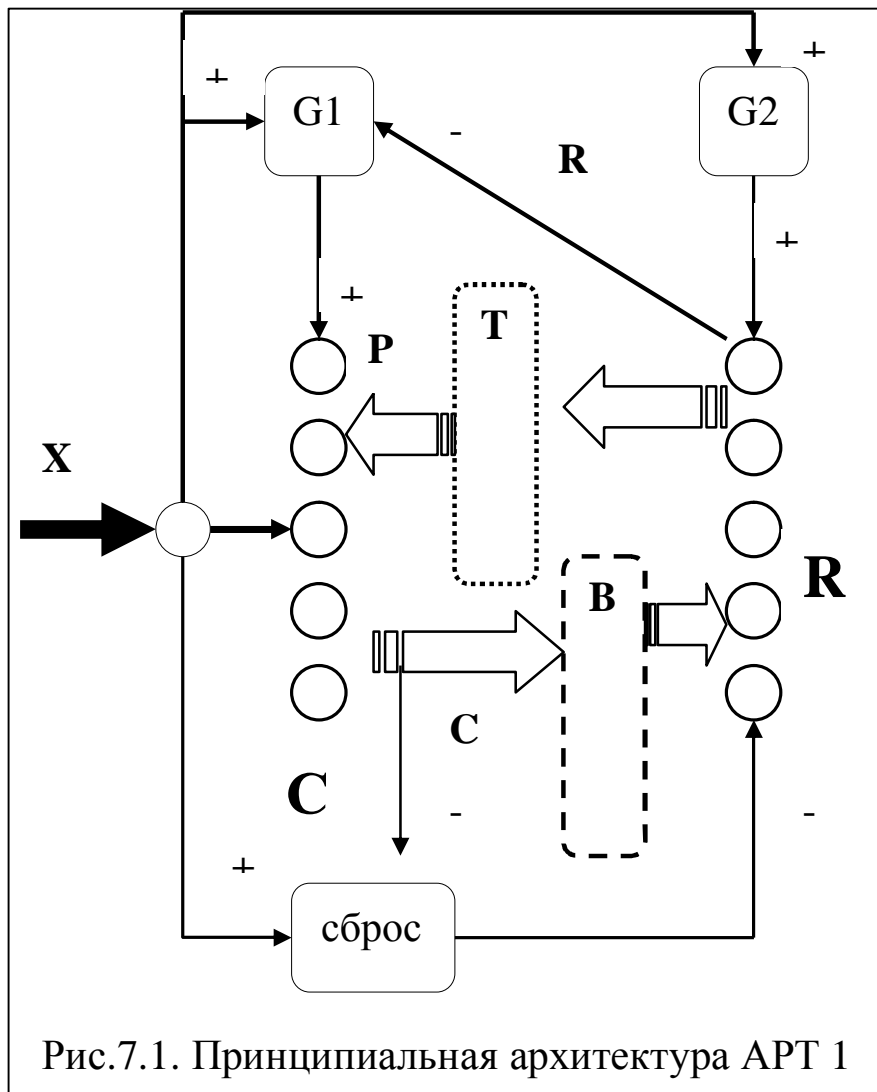


Рис.7.1. Принципиальная архитектура ART 1

Принципиальная архитектура ART1 (Рис.7.1) может быть представлена в виде сети из двух нейронных слоев – сравнения C и распознавания R , соединенных матрицами памяти B и T , двух приемников $G1$ и $G2$ и блока сброса.

Слой сравнения. Входной вектор X поступает на слой сравнения C . На первом этапе слой C пропускает его неизменным для формирования выходного вектора S . Этот вектор S затем будет модифицирован вектором P поступившим из слоя распознавания.

Каждый нейрон в слое сравнения (Рис. 7.2) получает три двоичных входа (0 или 1): (1) компонента x_i входного вектора X ; (2) сигнал обратной связи R_i как взвешенная сумма выходов распознающего слоя; (3) вход от $G1$ (один и тот же управляющий сигнал подается на все нейроны слоя C).

В соответствии с правилом 2/3 на выходе нейрона формируется единичное значение если как минимум два из трех его входов получают единицу; в противном случае на выходе нейрона формируется нуль. Поэтому изначально выходной сигнал $G1$ установлен в единицу, а все

компоненты вектора \mathbf{R} установлены в 0. Тем самым на первом этапе вектор \mathbf{C} идентичен входному вектору \mathbf{X} .

Слой распознавания R классифицирует входные вектора посредством реализации идеологии WTA – каждому запомненному образу назначается свой нейрон с вектором весов, соответствующим весам входного вектора. Таким образом, слой распознавания работает как слой Кохонена, рассмотренный в предыдущей лекции. Вектор весов описывается действительными числами. Двоичная версия образа одновременно запоминается в наборе весов слоя сравнения - этот набор состоит из весов связей, соединяющих определенные нейроны слоя распознавания, один вес на каждый нейрон слоя сравнения.

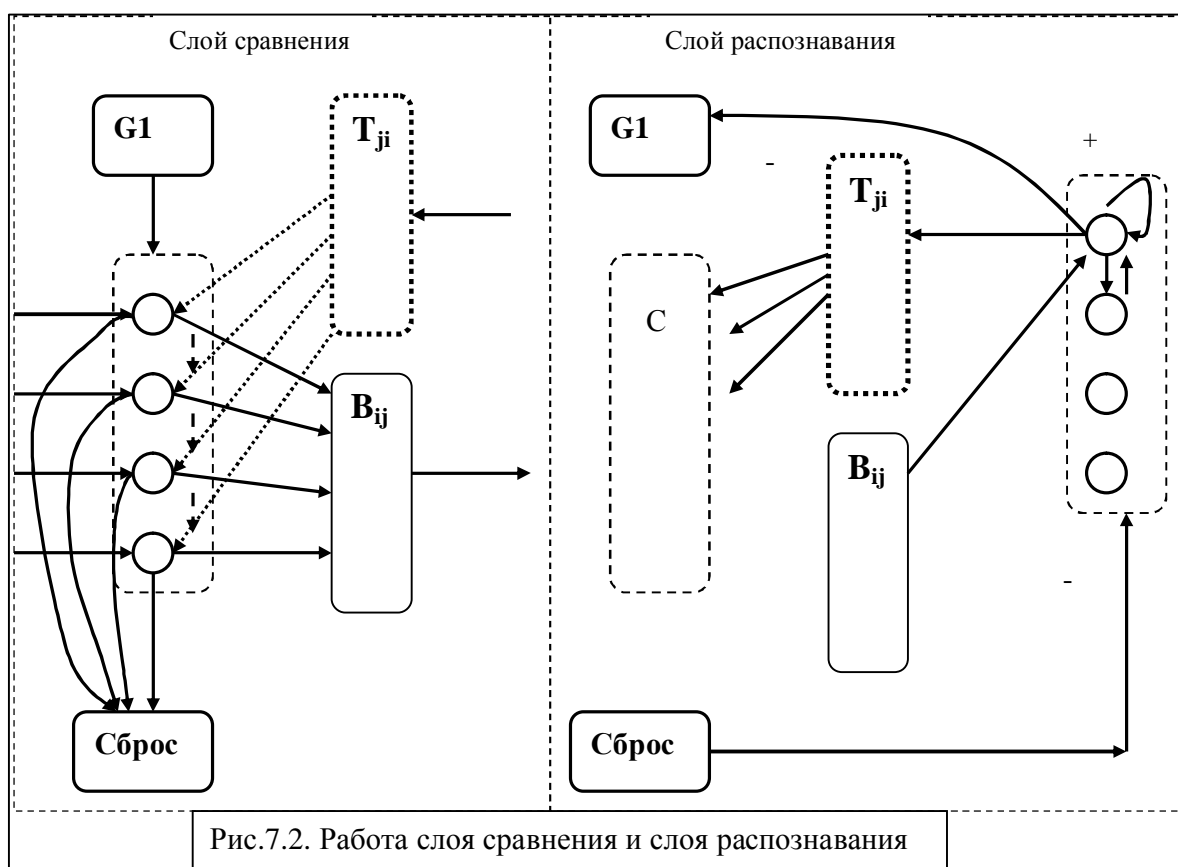


Рис.7.2. Работа слоя сравнения и слоя распознавания

Для реализации алгоритма WTA в слое R вводится *латеральное торможение* – каждый нейрон подает на другие нейроны отрицательный сигнал, пропорциональный его уровню возбуждения, и положительный сигнал, также пропорциональный уровню возбуждения, на самого себя. Тем самым нейрон с максимальным уровнем активации тормозит все остальные нейроны слоя и остается единственным возбужденным нейроном – победителем.

Приемники $G1$ и $G2$.

$G1$. Выходной сигнал $G1$ равен 1, если хотя бы одна компонента двоичного входного вектора \mathbf{X} равна единице; однако если хотя бы одна компонента вектора \mathbf{R} равна единице, $G1$ устанавливается в нуль.

G2. Выход G2, равен единице, если входной вектор \mathbf{X} имеет хотя бы одну единичную компоненту. Более точно, G2 является логическим ИЛИ от компонент вектора \mathbf{X} .

Модуль сброса. Вычисляет сходство между векторами \mathbf{X} и \mathbf{C} как отношение количества единиц в \mathbf{C} к их количеству в \mathbf{X} . Если это отношение ниже установленного порога, то вырабатывается сигнал сброса возбужденного нейрона в слое распознавания и сеть продолжает поиск среди ранее запомненных образов.

Этапы работы сети:

1. инициализация сети;
2. распознавание;
3. сравнение;
4. поиск;
5. обучение.

Этапы 2 – 4 относятся к фазе классификации.

1. Инициализация сети

Веса матрицы памяти \mathbf{V} устанавливаются $b_{ij} < \frac{L}{L-1+m}$ для всех i, j где m – размерность входного вектора, L – константа, большая 1 (обычно $L = 2$). Инициализация весов b_{ij} малыми значениями существенна для корректного функционирования АРТ. Если веса слишком большие, то ранее запомненный входной вектор будет активировать скорее свободный нейрон, чем ранее обученный. В результате все нейроны слоя распознавания окажутся связаны с одним входным образом. Если веса малы, то свободные нейроны не будут получать возбуждения большего,

чем ранее обученные нейроны в слое распознавания $b_{ij} = \frac{Lc_i}{L-1 + \sum_m c_i}$.

Веса \mathbf{T}_{ji} все инициализируются в единичные значения $t_{ji} = 1$ для всех i, j .

Параметр сходства ρ устанавливается в диапазоне от 0 до 1. Большие величины ρ ведут к высокой «разборчивости» сети, низкие – к слабой. В процессе функционирования величина ρ должна адаптивно изменяться.

2. Распознавание

При отсутствии входного вектора сигнал G2 устанавливается в 0 и, согласно правилу 2/3, все нейроны слоя распознавания имеют выход 0. Тем самым все нейроны слоя распознавания получают равные шансы на выигрыш в соревновании WTA.

Входной вектор \mathbf{X} имеет хотя бы одну компоненту, отличную от нуля, в результате на выходах G1 и G2 установятся единицы. Это позволяет нейронам слоя сравнения возбуждаться, если соответствующая

компонента входного вектора \mathbf{X} равна единице. Таким образом, на данном этапе вектор \mathbf{C} , выходящий из слоя сравнения, идентичен вектору \mathbf{X} .

В слое распознавания вычисляется свертка для каждого нейрона:

$$\text{NET}_j = (\mathbf{V}_j \cdot \mathbf{C}).$$

Нейрон с максимальным NET становится единственным победителем.

3. Сравнение

Нейрон-победитель возвращает единицу в слой сравнения через матрицу весов \mathbf{T} . Каждый весовой вектор T_j имеет двоичные значения весов, а каждый весовой вектор V_j представляет собой масштабированную версию соответствующего вектора T_j .

Поскольку вектор \mathbf{R} теперь не нулевой, то сигнал $G1$ устанавливается в нуль. Вектор \mathbf{P} суть *встречное ожидание*, т.е. тот из ранее запомненных векторов, который «как думает сеть» поступил на ее вход. Но сеть «может ошибаться». Поэтому производится сравнение \mathbf{P} и \mathbf{X} - в соответствии с правилом двух третей, возбуждаться будут только нейроны, получившие на входе одновременно единицы как от входного вектора \mathbf{X} , так и от возвращенного вектора \mathbf{P} .

Если вектор \mathbf{P} похож на вектор \mathbf{X} , то в векторе \mathbf{C} будут в основном единицы – классификация прошла успешно.

Если вектор \mathbf{P} отличен от вектора \mathbf{X} , то в векторе \mathbf{C} будут в основном нули и по критерию $S=N/D < \rho$, где D – количество единиц в векторе \mathbf{X} . N – количество единиц в векторе \mathbf{C} блок сброса выработает сигнал сброса, который затормозит возбужденный нейрон-победитель с слое \mathbf{R} и сеть продолжит поиск наиболее подходящего из ранее запомненных образов.

4. Поиск. Торможение возбужденного на предыдущем этапе нейрона-победителя в слое распознавания ведет к обнулению всех компонент вектора \mathbf{R} , $G1$ устанавливается в 1 и входной вектор \mathbf{X} еще раз предъявляется в качестве \mathbf{C} . Теперь только другой нейрон может стать победителем в слое распознавания, соответственно, и другое встречное ожидание \mathbf{P} будет предъявлено слою сравнения. Если и на этот раз \mathbf{P} не соответствует \mathbf{X} , то и этот нейрон в слое распознавания будет заторможен. Процесс поиска среди ранее запомненных образов будет продолжен до тех пор, пока не наступит один из двух исходов:

а. Будет найден ранее запомненный образ, сходство которого с входным вектором \mathbf{X} превышает порог, т.е. $S > \rho$. В этом случае процесс классификации остановится и сеть перейдет в режим дообучения для модификации матриц весов \mathbf{T}_j и \mathbf{V}_j , связанных с нейроном-победителем в слое распознавания.

б. Среди запомненных образов не будет найдено ни одного, соответствующего входному. В этом случае свободный нейрон в слое распознавания, веса которого равны единице, будет назначен этому образу. В соответствии с правилом 2/3 вектор \mathbf{C} будет идентичен входному вектору \mathbf{X} , $S=1$ и условие $S > \rho$ будет выполнено. Матрицы весов \mathbf{V}_j и \mathbf{T}_j будут настроены для соответствия новому входному образу.

5. Обучение. Веса матрицы B определяются по следующему правилу:

$$b_{ij} = \frac{Lc_i}{L-1 + \sum_m c_i},$$

где c_i – i -я компонента выходного вектора слоя сравнения; j – номер выигравшего нейрона в слое распознавания; b_{ij} – вес связи, соединяющей нейрон i в слое сравнения с нейроном j в слое распознавания; L – константа > 1 (обычно 2). Сумма в знаменателе представляет собой количество единиц на выходе слоя сравнения, т.е. «размер» этого вектора. В такой интерпретации «большие» вектора C производят меньшие величины весов b_{ij} , чем «маленькие» вектора C . Это свойство самомасштабирования делает возможным разделение двух векторов в случае, когда один вектор является поднабором другого; т. е. когда набор единичных компонент одного вектора составляет подмножество единичных компонент другого.

Для понимания важности свойства самомасштабирования рассмотрим пример классификации ранее известных сети векторов $X_1=10000$ и $X_2=11100$.

Пусть $L=1$. Тогда веса матрицы связей определяются следующим образом:

$$\left\{ \begin{array}{l} b_1^1 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1 \cdot 1}{1-1+1} = 1 \\ b_1^2 = b_1^3 = b_1^4 = b_1^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1 \cdot 0}{1-1+1} = 0 \end{array} \right. \Rightarrow B_1 = [1 \ 0 \ 0 \ 0 \ 0]$$

$$\left\{ \begin{array}{l} b_2^1 = b_2^2 = b_2^3 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1 \cdot 1}{1-1+3} = 1/3 \\ b_2^4 = b_2^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1 \cdot 0}{1-1+3} = 0 \end{array} \right. \Rightarrow B_2 = \left[\begin{array}{ccccc} 1/3 & 1/3 & 1/3 & 0 & 0 \end{array} \right].$$

Нетрудно видеть, что если масштабирования нет, то при предъявлении сети вектора $X_2=11100$ оба нейрона в слое распознавания будут иметь одинаковый уровень активации, равный 1.

Если $L=2$, то $B_1 = 10000$ и $B_2 = 1/2 \ 1/2 \ 1/2 \ 0 \ 0$, тогда, подавая на вход сети вектор X_1 , получим возбуждение 1,0 для нейрона 1 в слое распознавания и $1/2$ для нейрона 2; таким образом, нейрон 1 (правильный) станет победителем. Аналогично, предъявление вектора X_2 возбудит нейрон 1 до уровня 1, и нейрон 2 до уровня $3/2$, тем самым снова правильно выбрав победителя.

Покажем, что выбор значения параметра L позволяет изменять логику работы слоя распознавания. Рассмотрим случай сети, обученной векторами $X_1=10000$ и $X_2=11100$. Пусть на вход поступает вектор $X_3 = 1 \ 1 \ 0 \ 0 \ 0$.

Если $L=1$, то возбуждение нейрона 1 в слое R будет 1,0, а нейрона 2 – $2/3$. Нейрон 1 побеждает. Сеть считает, что вектор $X_3 = 1 \ 1 \ 0 \ 0 \ 0$ более похож на вектор $X_1=10000$, чем на вектор $X_2=11100$.

Если $L=1.5$, то:

$$\left\{ \begin{array}{l} b_2^1 = b_2^2 = b_2^3 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1.5 \cdot 1}{1.5-1+3} = \frac{3}{7} \\ b_2^4 = b_2^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1.5 \cdot 0}{1.5-1+3} = 0 \end{array} \right. \Rightarrow B_2 = \left[\frac{3}{7} \quad \frac{3}{7} \quad \frac{3}{7} \quad 0 \quad 0 \right],$$

$$\left\{ \begin{array}{l} b_1^1 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1.5 \cdot 1}{1.5-1+1} = 1 \\ b_1^2 = b_1^3 = b_1^4 = b_1^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{1.5 \cdot 0}{1.5-1+1} = 0 \end{array} \right. \Rightarrow B_1 = [1 \quad 0 \quad 0 \quad 0 \quad 0].$$

И при предъявлении вектора $\mathbf{X}_3 = 1 \ 1 \ 0 \ 0 \ 0$ возбуждение нейрона 1 в слое R будет 1,0, а нейрона 2 - 6/7. Нейрон 1 побеждает. Сеть считает, что вектор $\mathbf{X}_3 = 1 \ 1 \ 0 \ 0 \ 0$ более похож на вектор $\mathbf{X}_1=10000$, чем на $\mathbf{X}_2=11100$.

Если $L=2$, то возбуждение нейрона 1 в слое R будет 1,0, а нейрона 2 – тоже 1. Сеть считает, что вектор $\mathbf{X}_3 = 1 \ 1 \ 0 \ 0 \ 0$ одинаково похож как на вектор $\mathbf{X}_1=10000$, так и на вектор $\mathbf{X}_2=11100$. Какой нейрон в этом случае победит – дело случая, на решение могут повлиять малейшие флуктуации.

Если $L=3$, то матрица весов определяется:

$$\left\{ \begin{array}{l} b_1^1 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{3 \cdot 1}{3-1+1} = 1 \\ b_1^2 = b_1^3 = b_1^4 = b_1^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{3 \cdot 0}{3-1+1} = 0 \end{array} \right. \Rightarrow B_1 = [1 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$\left\{ \begin{array}{l} b_2^1 = b_2^2 = b_2^3 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{3 \cdot 1}{3-1+3} = \frac{3}{5} \\ b_2^4 = b_2^5 = \frac{L \cdot c_i}{L-1 + \sum_m c_i} = \frac{3 \cdot 0}{3-1+3} = 0 \end{array} \right. \Rightarrow B_2 = \left[\frac{3}{5} \quad \frac{3}{5} \quad \frac{3}{5} \quad 0 \quad 0 \right]$$

и возбуждение нейрона 1 в слое R будет 1,0, а нейрона 2 – 6/5. Нейрон 2 побеждает. Обученная таким образом сеть считает, что вектор $\mathbf{X}_3 = 1 \ 1 \ 0 \ 0 \ 0$ более похож на вектор $\mathbf{X}_2=11100$, чем на вектор $\mathbf{X}_1=10000$.

Тем самым, меняя значение параметра L , мы меняем логику работы сети – ее дискриминационную способность, т.е. способность различать (или не различать) близкие вектора.

Для правильного функционирования сети необходима инициализация весов матрицы \mathbf{B} малыми значениями, так при инициализации большими значениями возможен вариант, когда ранее запомненный входной вектор будет активировать не «свой» нейрон, а новый – сеть будет лишена способности «вспоминать», бесконечно обучаясь и ничего не вспоминая.

Веса матрицы \mathbf{T}_j , связанной с новым запомненным вектором, изменяются так, чтобы они стали равны соответствующим двоичным

величинам вектора C : $t_{ij} = c_i$ для всех i , где t_{ij} – вес связи между нейроном-победителем j в слое распознавания и нейроном i в слое C .

Поиск. Пусть на вход поступает вектор $X_3 = 11000$. Возбуждение нейрона 1 в слое R будет 1,0, а нейрона 2 - 2/3. Нейрон 1 побеждает, несмотря на то, что он не лучшим образом соответствует входному вектору, вектор C становится равным 11000 , S будет равно 1/2. Если уровень сходства установлен в 3/4, то нейрон 1 будет заторможен и победителем в следующем цикле распознавания станет нейрон 2. C примет значение 11000 , следовательно, S примет значение 1, и в силу выполнения критерия сходства поиск будет остановлен.

Можно видеть, что сеть ART не в полной мере аналогична биологическим прототипам. Например, известно, что биологическая память обладает свойством распределенности – повреждение отдельных нейронов не уничтожает запомненный образ, хотя и ухудшает характеристики его воспроизведения. В сети ART каждый образ связан с единственным нейроном и его повреждение ведет к исчезновению образа.

Для развития концепции ART важны теоретические результаты, полученные Гроссбергом. Наиболее значимы из них следующие:

1. В процессе поиска сеть сходится к устойчивому состоянию. После определения выигравшего нейрона в слое распознавания другие нейроны слоя распознавания возбуждаться не будут.
2. Процесс обучения сходится к устойчивому состоянию. Обучение не вызывает переключения с одного нейрона-победителя на другой.
3. Процесс обучения конечен. После конечного количества эпох для любой последовательности произвольных векторов достигаются стабильные состояния матриц весов. Повторяющиеся последовательности обучающих векторов не ведут к циклическому изменению весов.
4. После стабилизации процесса обучения предъявление одного из обучающих векторов (или вектора с существенными характеристиками категории) будет активизировать требуемый нейрон слоя распознавания без поиска. Такая характеристика "прямого доступа" обеспечивает быстрый доступ к предварительно изученным образам.

Оптические реализации ART. Идеология ART включает в себя как трудноформализуемые процедуры, характеризующиеся большим объемом параллельных вычислений (формирование матриц связей), так и строгологические. Соответственно, оптика применима в ART преимущественно для реализации первого класса процедур. Например, в работе [2] в качестве оптического процессора использован рассмотренный в предыдущих лекциях голографический коррелятор.

Литература к Лекции 7

1. Уоссермен Ф., Нейрокомпьютерная техника - М.: Мир, 1992.
2. D.C. Wunsch et al., "Photorefractive adaptive resonance neural network", Applied Optics, 1993, v.32, #8, p.1399-1407.

Миссия университета – генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

КАФЕДРА ФОТОНИКИ И ОПТОИНФОРМАТИКИ

<http://phoi.ifmo.ru>

Заведующий кафедрой – профессор, д.ф.-м.н., лауреат премии Ленинского комсомола, руководитель Российской научной школы «Фемтосекундная оптика и фемтотехнологии» Сергей Аркадьевич Козлов.

На кафедре российские и иностранные преподаватели готовят бакалавров и магистров по образовательному направлению 200700 «Фотоника и оптоинформатика» – одному из самых инновационных направлений современной науки и техники.

Уже начиная со второго курса, студенты кафедры имеют возможность практической работы в области фотоники и оптоинформатики, участвуя в исследованиях и разработках Международных научных центров и лабораторий кафедры, а также в ведущих фирмах Санкт-Петербурга. На кафедре разрабатываются и создаются:

- оптические и квантовые технологии сверхбыстрой передачи и записи информации;
- оптические системы искусственного интеллекта и сверхбыстродействующие оптические процессоры;
- информационные оптические системы, строящиеся на основе новых физических принципов, в том числе на нанотехнологиях.

Студенты кафедры участвуют в Международных научных конференциях, публикуют свои работы в ведущих мировых изданиях. Среди студентов и аспирантов кафедры есть стипендиаты Президента и Правительства Российской Федерации, победители конкурсов научных работ, проводимых Российской Академией наук, крупнейшими мировыми научными обществами, такими, как INTAS, SPIE, CRDF, OSA, IEEE Photonics Society.