

Реализация правдоподобных выводов на нейросетях со связями по схеме голографии Фурье¹

Аннотация. Показана возможность реализации абдуктивного и индуктивного выводов на двухслойной нейронной сети с двунаправленными связями, реализуемыми методом голографии Фурье. Определены требования к характеристикам паттернов внутренней репрезентации и активационным функциям нейронов. Приведены результаты численного моделирования.

Ключевые слова: правдоподобный вывод, абдукция, индукция, обработка образов, нейронные сети, голография.

Введение

Основной функцией интеллекта, как биологического, так и искусственного (ИИ), является рассуждение, под которым обычно понимают построение последовательности аргументов, вынуждающих принятие некоторого утверждения [1]. Традиционно принято выделять два типа рассуждений – достоверные, к которым относится дедуктивный вывод, и правдоподобные, к которым относится абдукция и индукция [1]. Основное внимание исследователей на протяжении веков, начиная, как минимум, с Аристотеля, привлекал достоверный вывод. Правдоподобные рассуждения также вызвали определенный, хотя и несколько меньший, интерес – им уделили внимание такие титаны мысли, как Аристотель и Ф.Бэкон. Однако в плане практических результатов, включая их реализацию методами ИИ, успехи изучения и моделирования правдоподобных рассуждений до недавнего времени были существенно скромнее, нежели достоверных [1,2]. При этом многие исследователи отмечали, что правдоподобные рассуждения являются не просто атрибутом человеческого интеллекта, но и важнейшей, неотъемлемой частью процесса познания,

позволяющей вывести познание за рамки непосредственно-чувственного восприятия. В настоящее время наблюдается явное возрастание интереса к правдоподобным рассуждениям, при этом большая часть усилий направлена на их моделирование на основе использования формально-логического аппарата, глубоко разработанного при исследованиях дедуктивного вывода, и реализацию рассуждений на универсальном компьютере.

Вместе с тем, известно, что даже задачи, адекватно описываемые в рамках логико-алгебраического формализма, биологический мозг решает отнюдь не так же, как универсальный компьютер, а посредством обработки образов на нейросетевых структурах [3]. Эти образы суть не философские абстракции, но физически реальные картины нейронной активности коры мозга – паттерны внутренней репрезентации (ПВР) воспринимаемой или вспоминаемой информации. Результат рассуждения также представляет собой ПВР. Причем в ряде случаев мозг, «безнадёжно медленный по сравнению с компьютером», решает задачи не просто эффективнее компьютера, но решает и те задачи, справиться с которыми логико-алгебраическими методами не удастся [4,5]. При этом по-

¹ Работа выполнена при поддержке РФФИ, проект 09-01-00165-а.

пытки моделирования образного мышления в рамках компьютерной парадигмы встречаются ряд трудностей. Поэтому ряд исследователей обращает внимание на другие, прежде всего, – нейросетевую парадигму [3].

Известно также, что противопоставление логического и образного мышления, основанное на реальной функциональной асимметрии полушарий мозга [6], носит в определенной степени искусственный характер, поскольку в реальном процессе рассуждений, реализуемом биологическим мозгом, эти две формы неразрывны [4, 5, 7]. Поэтому представляет интерес возможность реализации правдоподобных рассуждений на искусственных нейронных сетях, т.е. в рамках концепции образного мышления.

Одно из направлений в ИИ, акцентирующее внимание на образной форме мышления, – голографическая парадигма, истоки которой можно проследить еще с работ Гольдшайдера 1906 года [8]. Голографическая парадигма может рассматриваться также как составная часть нейросетевой парадигмы в силу того, что голография как технология записи и восстановления комплексных волновых полей полностью удовлетворяет основным атрибутам нейросетевой парадигмы – принципам обучаемости, связности, распределенности, ассоциативности [3].

Вместе с тем, большая часть работ в рамках голографической парадигмы посвящена реализации ассоциативной памяти. Память, безусловно, является фундаментом интеллекта, но сама по себе еще не интеллект. Поэтому представляет интерес реализация рассуждений в рамках голографической парадигмы.

В работах [9, 10] было показано, что схема голографии Фурье (ФГ), адекватно описываемая моделью двухслойной нейронной сети, строит алгебру нечеткозначимой логики с Фурье-дуальными определяющими моделью операциями и экспериментально реализовано правило вывода «Обобщенный Modus Ponens». В работе [11] предложен и экспериментально подтвержден подход к интеграции логической (в классе нечеткозначимых логик) и образной форм мышления методом ФГ. В работе [12] показана реализация методом ФГ логики с исключениями. В настоящей работе, в развитие этого подхода, предложен подход к реализации абдуктивного и индуктивного выводов на двухслойной нейронной сети, адекватной реальной схеме голографии Фурье.

1. Биологически мотивированный подход и нейросетевая модель

1.1. Подход к задаче

Для большей наглядности используем в дальнейших построениях классический пример классификации рассуждений (по Ч. Пирсу) [1]:

Дедукция: Правило (*Все люди смертны*) (1)
 Случай (*Сократ человек*) (2)
 Результат (*Сократ смертен*) (3)

Индукция: (2) Случай (*Сократ человек*)
 (3) Результат (*Сократ смертен*)
 (1) Правило (*Все люди смертны*)

Абдукция: (1) Правило (*Все люди смертны*)
 (3) Результат (*Сократ смертен*)
 (2) Случай (*Сократ человек*)

Посмотрим, как эти рассуждения могут быть реализованы методами искусственных нейронных сетей (НС). Для этого, прежде всего следует определить методы представления посредством паттернов (ПВР) операндов, входящих в вышеприведенные связки, а затем – характеристики этих ПВР, обрабатываемых НС, существенные для определения технически конкретных методов их ассоциирования.

В приведенных схемах рассуждений используются всего три типа операндов и, соответственно, паттернов:

- ПВР «человек», представляющий свойство быть человеком вообще (а не обезьяной, например), который обозначим Im_H ,
- ПВР, представляющий свойство смертности вообще, абстрагируясь от частных, который обозначим Im_D ,
- ПВР, представляющие только индивидуальные признаки персоны (Сократа, Платона, Пупкина, etc.), которые будем обозначать Im_{Ind} с указанием при необходимости имени персоны или свойства.

Нетрудно видеть, что два первых ПВР можно определить скорее как понятия, а последний – как образ. Напомним, что в рамках концепции образного мышления принято выделять (и отчасти противопоставлять) две категории – образ и понятие [6, 13-14]. При этом как образ, так и понятие суть ПВР, реально (физически) формирующиеся в коре головного мозга. Различие между ними в том, что образ относится к категории сенсорных ПВР, т.е. формируемых на

основе сенсорной информации (воспринимаемой, а затем вспоминаемой), а понятие – это ПВР, формируемый мозгом в результате обработки сенсорной информации. Образ всегда характеризуется определенной целостностью, это гештальт [14], понятие же формируется в процессе разрушения гештальта, вычленения из него только существенных для понятия признаков и абстрагирования от несущественных, но неотъемлемых признаков для целостного образа. При этом условная шкала «образ – понятие» является отнюдь не бинарной, но непрерывной – переход от образа к понятию происходит постепенно, большая часть хранящихся в памяти паттернов включает в себя как сенсорную, так и вербальную информацию, т.е. находится где-то в интервале между сугубо сенсорным образом и «чистым» понятием. В частности, в использованном примере ПВР Im_H , отнесенный нами к категории «скорее понятия», включает в себя и сенсорную составляющую. Некоторые аспекты формирования понятия из сенсорного образа подробнее рассмотрим ниже, в части, посвященной реализации индукции.

Таким образом, правило (1) связывает два понятия - Im_H и Im_D .

Случай (2) связывает ПВР - образ, представляющий только индивидуальные свойства персоны, который обозначим $Im_{IndPerson}$, с понятием Im_H . Здесь важно, что по условиям примера ПВР $Im_{IndPerson}$ (сугубо сенсорный образ), не включает в себя ПВР Im_H (понятие), поскольку о том, что данная персона - человек, мы узнаем только в результате связи двух этих ПВР. Установление этой связи и является целью интеллектуальной деятельности. В рамках приведенного примера такая схема выглядит несколько надуманной, но для других, более сложных, примеров, относящихся к методологии научного познания [1], она вполне реальна. Кроме того, этот простой пример нужен нам для понимания излагаемого подхода к решению поставленной задачи.

Результат (3) так же, как и случай (2), связывает сенсорный образ $Im_{IndPerson}$ с понятием Im_D .

Рассмотрим возможность реализации этих связок двухслойной нейронной сетью (НС) с двунаправленными связями, приведенной на Рис.1. Выбор этой простой архитектуры обусловлен тем, что она адекватна реальной схеме голографии Фурье [9-12]. Не вдаваясь в описание схемы голографии Фурье, отметим лишь,

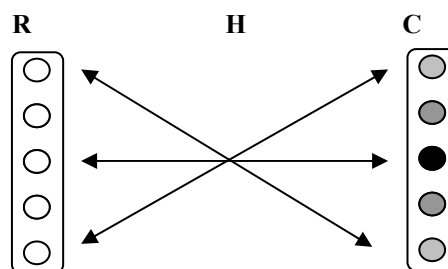


Рис.1. Модель двухслойной нейронной сети с двунаправленными связями, адекватная схеме голографии Фурье

что особенностью данной модели НС является реализация весов межнейронных связей в пространстве Фурье в виде

$$H_{RC}(v_x, v_y) = F^*(Im_R(x, y))F(Im_C(\Delta_x, \Delta_y)), \quad (4)$$

где F – символ преобразования Фурье, $Im_R(x, y)$ и $Im_C(\Delta_x, \Delta_y)$ - паттерны, описывающие состояние нейронных слоев R и C на этапе обучения, астериск обозначает комплексное сопряжение, а (v_x, v_y) – частоты (при обработке изображений как двумерных функций, определенных на плоскости, - пространственные частоты). В дальнейшем изложении для упрощения выражений аргументы опустим. Выражение (4) описывает голограмму Фурье. Нетрудно видеть, что данная модель, как и голографическая схема, реализует правило обучения Хэбба.

	Слой R	Слой C
Правило (1)	Im_H	Im_D
Результат (3)	Im_{Ind}	Im_D

В таблице приведен вариант представления ПВР слоями НС, позволяющий реализовать как правило (1), так и результат (3). Для того, чтобы реализовать в этой НС рассуждения, включающие все три связки (1-3), необходимо в ней же ассоциировать паттерны Im_H и Im_{Ind} применительно к случаю (2). Очевидно, что попытка представить эти ПВР каждый в своем нейронном слое с установлением между ними матрицы связей, подобно тому, как это сделано для связок (1) и (2), приведет к выходу за рамки архитектуры Рис.1 – добавлению еще одного, третьего, слоя. Следовательно, необходимо связать Im_H и Im_{Ind} так, чтобы все три связки (1-3) реализовывались двухслойной сетью параллельно, без изменения архитектуры и переобучения сети. Для ответа на вопрос о методе свя-

зи паттернов Im_H и Im_{Ind} рассмотрим сначала реализацию в данной НС дедуктивного вывода, а затем перейдем к правдоподобным выводам - абдукции и индукции.

1.2. Выбор метода представления информации на примере дедуктивного вывода

1.2.1. Связь паттерна индивидуальных признаков с паттерном «человек». Шкала общности свойств

Для реализации дедуктивного вывода сеть должна быть обучена правилу (1)

$$Im_H \leftrightarrow Im_D$$

и, при предъявлении ПВР $Im_{IndPerson}$, реализовать цепочку ассоциаций

$$Im_{IndPerson} \rightarrow Im_H \rightarrow Im_D,$$

где первая ассоциация представляет случай (2). Иными словами, НС должна некоторым образом распознавать, что и Сократу, и Платону, и Пупкину присуще общее для всех них понятие «человек».

Поскольку классический нейросетевой вариант реализации связки $Im_{IndPerson} \rightarrow Im_H$ выше был нами отвергнут, то в рамках принятого подхода (Табл. 1, Рис.1.) рассмотрим другой метод, основанный на том хорошо известном факте, что ПВР каждой персоны представляет признаки, различающиеся по степени своей общности – от глобальных до локальных.

Интерпретация этого тезиса достаточно очевидна в случае визуального распознавания - на большом расстоянии мы можем не узнать, кого именно видим, но понимаем, что это человек. Ближе – различаем мужчину это или женщина, еще ближе – оцениваем другие признаки, ближе – узнаем конкретную персону. Таким образом, может быть введена шкала общности свойств или признаков. В качестве такой шкалы используем шкалу частот в Фурье-пространстве, в котором формируется матрица связей нейронных слоев H в виде голограммы Фурье (4). Тогда наиболее общие свойства (быть человеком) соответствуют самым низким частотам в спектре ПВР, далее, по мере увеличения частоты возрастает и индивидуализация свойств.

Здесь следует особо подчеркнуть условность использованной аналогии, поскольку ПВР представляет собой совсем не то изображение, что формирует оптическая система (глаз, например), но результат преобразования этого

изображения сенсором и сенсорным трактом. ПВР, как картина распределения нейронной активности, совершенно не похож на изображение. В этой связи отметим, что выбор шкалы частот в качестве шкалы общности свойств имеет своим обоснованием не столько вышеприведенную интуитивно понятную интерпретацию из области зрительного восприятия, сколько то, что ПВР, представляющие общие свойства, должны обладать большей внутренней связностью, т.е. иметь больший радиус корреляции. Именно большой радиус корреляции обеспечивает большую устойчивость ПВР общих свойств против локальных повреждений или искажений. Действительно, критичность амплитуды глобального максимума автокорреляционной функции к искажениям объектного изображения относительно эталонного (для широкого класса искажений) в общем случае определяется параметром

$$\frac{L_x L_y}{r_x r_y},$$

где L_x, L_y – геометрические размеры изображения, r_x, r_y – радиус корреляции [15, 16]. А больший радиус корреляции определяется именно большим удельным весом низких частот в спектре [17, 18].

ПВР, представляющий индивидуальные свойства, напротив, может иметь очень малый радиус корреляции, чем обеспечивается богатство индивидуальных свойств личности, в том числе, казалось бы взаимоисключающих. В результате того, что ПВР персоны объединяет субпаттерны как с высокой внутренней коррелированностью (низкочастотные), так и с низкой (высокочастотные), достигается и гармоничность личности как сочетание целостности с её индивидуальным богатством, и способность развиваться, оставаясь при этом самой собой.

Таким образом, ПВР персоны может быть представлен в виде суммы субпаттернов, каждый из которых репрезентирует признаки, отличающиеся от признаков другого субпаттерна степенью своей общности. Крайние значения шкалы представлены субпаттернами Im_H и Im_{Ind} , а между ними на шкале общности расположены другие субпаттерны, представляющие иные свойства и признаки персоны (пол Im_G , возраст Im_A , etc.). Эти субпаттерны накладываются друг на друга в пространстве паттернов, имея одну область определения (в слое R), но

разнесены в пространстве Фурье, поскольку занимают разные частотные диапазоны, как это схематически изображено на Рис.2.).

$$Im_{Person} = Im_H + \dots + Im_{Ind} \quad (5)$$

1.2.2. Характеристики паттерна «смертен» при реализации дедукции

Обучение сети Рис.1 правилу (1) заключается в формировании матрицы связей H_{HD} в соответствии с (4), связывающей низкочастотный субпаттерн Im_H с паттерном, представляющим в слое С свойство «смертен» Im_D .

$$H_{HD} = F^*(Im_H)F(Im_D) \quad (6)$$

В пространстве Фурье частотный диапазон паттерна Im_D должен пересекаться с частотным диапазоном субпаттерна Im_H - в противном случае матрица связей (6) сформирована быть не может. Поскольку выше мы определили, что субпаттерн Im_H низкочастотный, то и паттерн Im_D при реализации дедуктивного вывода также должен быть низкочастотным (или включать в себя низкочастотную спектральную компоненту). Наличие или отсутствие в паттерне Im_D высокочастотной компоненты на результат дедуктивного вывода влияет только с точки зрения величины отношения сигнал/помеха.

Случай (2) реализуется предъявлением обученной сети паттерна персоны (5), включающего в себя субпаттерн Im_H , в результате сеть формирует отклик

$$\begin{aligned} Out &= F [F(Im_{Person})F^*(Im_H)F(Im_D)] = \\ &= F [F(Im_H)F^*(Im_H)F(Im_D) + \\ &\quad + F(Im_{Ind})F^*(Im_H)F(Im_D)] = \\ &= (Im_H \otimes Im_H)^* Im_D = Im_D \end{aligned}$$

где символы \otimes и $*$ обозначают свертку и корреляцию, соответственно, а в предположении отсутствия высокочастотной спектральной компоненты в составе Im_D $F(Im_{Ind})F^*(Im_H) = 0$. Таким образом, получаем требуемый результат (3) – «данная персона смертна».

Если в составе Im_D высокочастотная компонента присутствует, то, поскольку о смертности данной персоны нам по условиям задачи ничего не известно, компонента корреляционного поля $F(F(Im_{Ind})F^*(Im_H))$ является помехой, а корреляционный алгоритм, как известно, осуществляет оптимальное по критерию отношения сигнал-помеха выделение сигнала из шума, т.е. также получаем требуемый результат

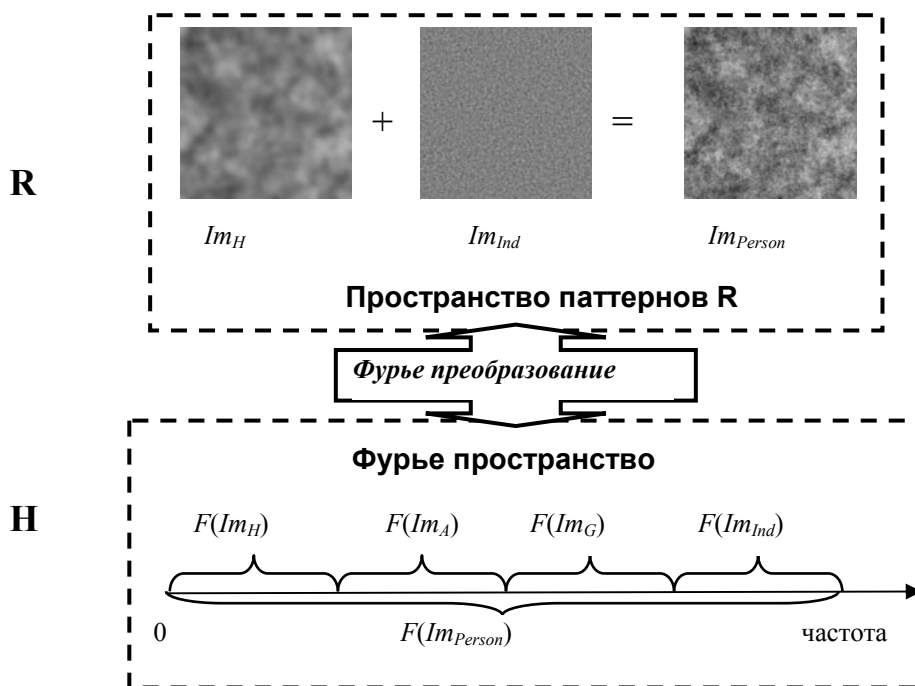


Рис.2. Схема представления случая (2) «Данная персона – человек» В слое R нейронной сети Рис.1 и пространстве Фурье H

Субпаттерн Im_{Ind} представляет индивидуальные свойства персоны, а субпаттерн Im_H – свойство «человек».

«данная персона смертна», но с несколько меньшим отношением сигнал – помеха [15]

$$\begin{aligned}
 Out &= F [F(Im_{Person}) F^*(Im_H) F(Im_D)] = \\
 &= F [F(Im_H) F^*(Im_H) F(Im_D)] + \\
 &\quad + F [F(Im_{Ind}) F^*(Im_H) F(Im_D)] = Im_D + N,
 \end{aligned}$$

где N обозначает кросс-корреляцию Im_{Ind} и Im_D , т.е. помеху.

1.3. Реализация индуктивного вывода

При реализации индуктивного вывода сеть должна обучаться примерами результатов (3), связывающих паттерн персоны (5) с ПВР свойства «смертен», которое рассматривается здесь

(по условиям задачи индукции) как сугубо индивидуальное – обозначим его Im_{DInd} . Матрица связей, формируемых в результате обучения, описывается выражением

$$\begin{aligned}
 H_{PDInd} &= F^*(Im_{Person}) F(Im_{DInd}) = F^*(Im_H) F(Im_{DInd}) + \\
 &\quad + F^*(Im_{Ind}) F(Im_{DInd}) = F^*(Im_{Ind}) F(Im_{DInd}) \quad (7)
 \end{aligned}$$

Если при реализации дедукции частотный диапазон Im_D пересекался с диапазоном Im_H , то есть Im_D был низкочастотным, то теперь $F(Im_{DInd})$ должен пересекаться с $F(Im_{Ind})$. По условиям задачи пересечения $F(Im_{DInd})$ с $F(Im_H)$ на этапе обучения нет - $F(Im_{DInd})$ локализован только в области высоких частот (Рис.3). Задача

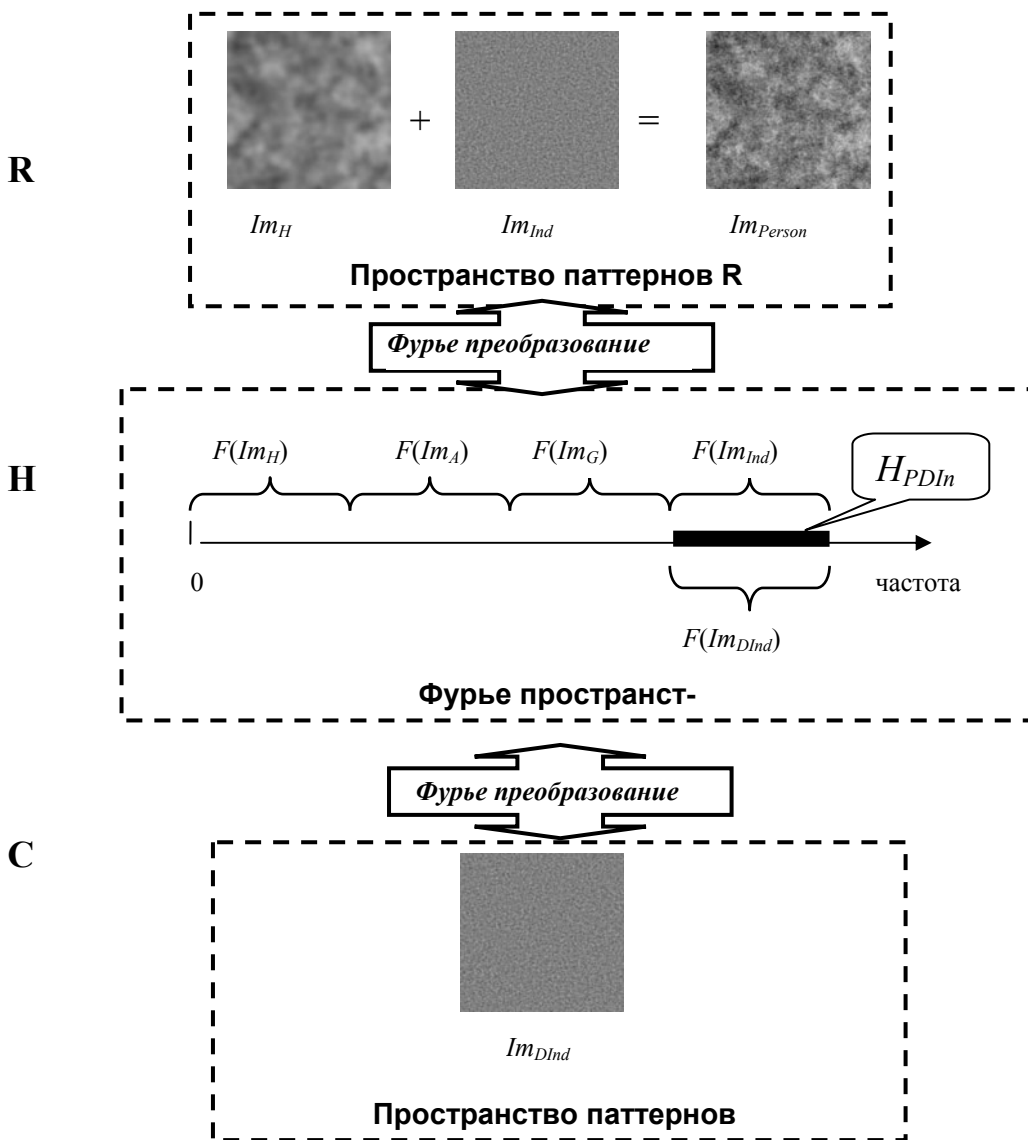


Рис.3. Схема обучения сети результатом (3) «Данная персона смертна» - матрица связей (голограмма) (7) в Фурье-пространстве формируется только в области высокочастотного субпаттерна $F(Im_{DInd})$

индукции заключается, таким образом, в терминах нашего подхода в расширении спектра Im_{DInd} в область низких частот с последующей записью голограммы, связывающей новый ПВР с расширенным спектром, который обозначим Im_{DExt} , с Im_H .

Никаких других требований на Im_{DExt} , кроме наличия низкочастотных компонент в спектре, для ясности изложения основной идеи сейчас не накладываемся. Частный случай решения этой задачи в пространстве паттернов - преобразование произвольного паттерна в дельта функцию, поскольку спектр последней имеет постоянную спектральную плотность. С технической точки зрения, эта задача, как задача реализации хорошо известной в теории НС идеологии WTA (победитель забирает все), очень просто решается использованием в слое С нелинейного итерирующего отображения – за несколько итераций из сложного сигнала или изображения выделяется единственный максимум, а остальные сводятся к нулю. Соответственно, и спектр с каждой итерацией расширяется, в пределе - в спектр с постоянной спектральной плотностью. Результат каждой итерации, как единичный интеллектуальный акт, должен формировать новую матрицу связей – перезаписывать голограмму.

Итерационный метод имеет очевидные биологические основания – понимание того, что «все люди смертны» приходит в результате раздумий над фактами смертности отдельных людей, или даже одной персоны. Эти раздумья в рамках схемы Рис.1 выражаются в многократной активации нейронов слоя С, в котором представлен Im_{DInd} , а многократность активации нейронов обеспечивается многократным прохождением возбуждения от слоя R к слою С и обратно с перезаписью голограммы на каждой итерации. В терминах голографии, мы снова, как и при реализации дедукции, получили схему голографии Фурье, но с нелинейным обращением волнового фронта в плоскости корреляций.

Если активационная функция С-нейронов нелинейна (в голографии эта функция реализуется экспозиционной характеристикой регистрирующей среды), и эта нелинейность адекватна задаче выделения нейрона-победителя, то с каждой итерацией образ Im_{DInd} в слое С будет все более и более концентрироваться (обостряться), приближаясь к δ -функции.

$$Im_{DInd}^{k+1} = NI(Im_{DInd}^k)$$

$$\lim_{k \rightarrow \infty} (Im_{DInd}^k) = Im_{DExt} = \delta$$

где NI –нелинейная активационная функция нейронов слоя С, а k – номер итерации. Для этого достаточно, чтобы вторая производная нелинейной функции в пределах рабочего диапазона удовлетворяла условию

$$(NI(x))'' > 0.$$

Большинство реальных голографических регистрирующих сред имеют участок экспозиционной характеристики, удовлетворяющей этому условию – как правило, это диапазон недоэкспозиций и малых экспозиций [17, 18].

Подчеркнем важность не единственности результата (3), которым обучается НС. В принципе, индуктивное обобщение часто возможно и на единственном примере. Однако, мы не накладывали никаких ограничений на Im_{DInd} и, если Im_{DInd} окажется гармоническим колебанием, то итерационная процедура не позволит свести его к дельта-функции, какой бы вид нелинейности не использовался, - расширение спектра при этом возможно только в область более высоких частот. Но сложение с другим Im_{DInd} , пусть тоже гармоникой, но другой частоты, ведет к появлению комбинационных, в том числе, разностных, частот.

Подчеркнем также и другой принципиальный момент – Фурье-образ ПВР персоны, представленного в слое R, должен быть либо непрерывным, либо дискретным, но таким, чтобы пустые интервалы между спектральными компонентами были меньше, чем возможное за одну итерацию расширение спектра Im_{DInd} . В противном случае новый, расширенный за одну итерацию, спектр, обозначим его $F(Im_{DExt})$, будет по-прежнему пересекаться только с $F(Im_{Ind})$ и запишется опять та же самая матрица связей (7), что и первый раз. Размышляющий субъект должен знать не только данную персону как индивидуум, но и то, что данный индивидуум принадлежит к группе более общих классов. Иными словами, для того, чтобы раздумья были продуктивны, необходим как определенный запас знаний (заполненность спектрального диапазона), так и определенные интеллектуальные способности (степень нелинейности), чтобы результат единичного интеллектуального акта пересекался с внутренней картиной мира.

Соответственно, и индуктивная процедура расширяет знания последовательно – в нашем примере после первой итерации может прийти понимание того, что все представители данной возрастной группы смертны (голограмма H_{DExt1} на Рис.4.), после второй – все представители данного пола смертны (голограмма H_{DExt2} на Рис.4.) и т.д.

Способность к обобщению зависит как от активационной функции С-нейронов, так и от субъективной яркости образа смертности конкретных персон. Если нелинейность выражена

слабо, или имеет вид, не позволяющий выделить максимум, то обобщение не наступает. Здесь мы имеем еще одно косвенное подтверждение (с точки зрения сугубо физической) того, хорошо известного в когнитивной психологии, факта, что интеллектуальные способности человека определяются в значительной степени его способностью к реализации хаотической динамики, а последняя, как известно, определяется именно нелинейностью активационных функций его нейронов [19].

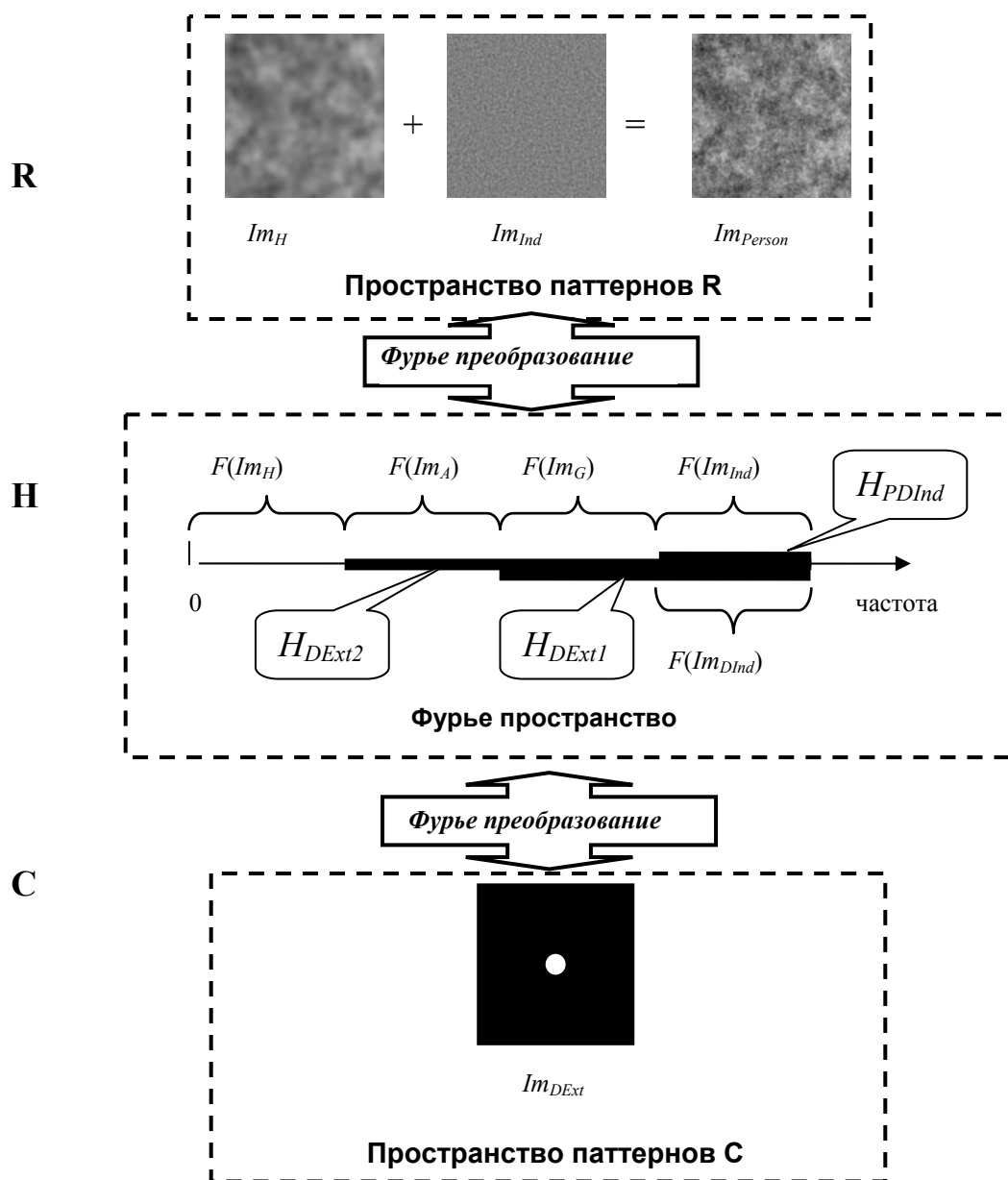


Рис.4. Схема реализации в сети итерационной процедуры – по мере трансформации в слое С образа Im_{DIn} в дельта-функцию, его спектр $F(Im_{DInd})$ расширяется, позволяя в результате каждой итерации записывать новые голограммы во все более широком частотном диапазоне - H_{DExt1} , H_{DExt2} , вплоть до H_{DExt} , перекрывающей весь частотный диапазон $F(Im_{Person})$.

Если свойство смертности неактуально для познающего субъекта, то и «яркость» Im_{DInd} невелика по сравнению с другими образами, одновременно с Im_{DInd} представленными в слое С – в результате итераций образ смертности не только не усиливается, но наоборот – ослабляется, а усиливаются другие, более актуальные для субъекта, образы и обобщение как понимание того, что люди смертны, приходит поздно, или не приходит вообще. Если же свойство воспринимается ярко, то уже нескольких итераций достаточно для того, чтобы пришло понимание - «все люди смертны».

Обратим внимание, что переход от сенсорного образа смертности конкретной персоны Im_{DInd} к понятию смертности вообще Im_{DExt} соответствует вышеупомянутому [14] отличию образа от понятия в плане наличия у первого целостности, т.е. внутренней коррелированности. Переход от произвольного паттерна Im_{DInd} к δ -функции и есть разрушение внутренней коррелированности, т.е. разрушение гештальта. Строго говоря, по условиям задачи, применительно к рассматриваемой реализации (рис.1.) требуется не переход к δ -функции, а переход в Фурье пространстве к постоянной спектральной плотности. Именно постоянная спектральная плотность обеспечивает общность понятия, его применимость к любому субпаттерну из совокупности (5), составляющей паттерн персоны.

Отметим, что введенное здесь понятие Im_{DExt} , характеризующееся дельта-коррелированностью, оказывается более общим, чем ранее введенное понятие Im_D , поскольку коррелирует как с сенсорными (высокочастотными) образами Im_{Ind} , так и с вербальными (низкочастотными) Im_H . В этом смысле оно всеобщее. Постоянная спектральная плотность, в свою очередь, через теорему Винера-Хинчина обуславливает дельта-коррелированность, вне зависимости от конкретного вида паттерна в слое С – это может быть как единичная δ -функция, так и δ -коррелированный паттерн.

Вместе с тем, переход в слое С от произвольного паттерна Im_{DInd} именно к единичной δ -функции обусловлен тем, что размерность области определения матрицы связей нейронных слоев равна сумме размерностей областей определения ассоциируемых паттернов и, в случае двух изображений, определенных на плоскости, матрица связей должна иметь 4-х мерную область определения, что физически

нереализуемо. В случае же единственной δ -функции в слое С, матрица связей определена на плоскости, т.е. физически реализуема в виде тонкой голограммы Фурье.

Таким образом, индуктивный вывод реализуется в рассмотренной архитектуре НС (Рис.1). Если для реализации дедукции было достаточно однократного прохождения от слоя R к слою С и, соответственно, однонаправленных связей, то для реализации индукции необходимы двунаправленные связи, позволяющие реализовать итерационную процедуру – не единичный интеллектуальный акт, а процедуру раздумий.

На Рис.5 приведен пример нелинейной активационной функции нейронов слоя С, на Рис.6 – результаты трансформации паттерна с ограниченным экспоненциальным спектром к дельта-функции, а на Рис.7 – соответствующие им действительные части Фурье-образов.

1.4. Реализация абдуктивного вывода

По условиям задачи имеется паттерн персоны Im_{Person} , содержащий только высокочастотную часть Im_{Ind} , представляющую индивидуальные признаки. Низкочастотный субпаттерн Im_H в Im_{Person} отсутствует. Задача абдукции в рамках нашего подхода – задача формирования в составе Im_{Person} низкочастотного субпаттерна Im_H .

Сеть (Рис.1) обучена правилу (1), связывающему в соответствии с (4) низкочастотный субпаттерн Im_H с Im_D , как это было сделано при реализации дедукции, т.е. матрица связей нейронных слоев имеет вид

$$H_{HD} = F^*(Im_H)F(Im_D). \quad (8)$$

Напомним, что (8) - низкочастотная голограмма, поскольку $F(Im_H)$ локализован только в области низких частот (Рис.2). При предъявлении сети, обученной таким образом, какого-либо паттерна в одном из слоев (R или С), отклик в другом слое (С или R, соответственно) будет сформирован только в том случае, если спектральный диапазон предъявляемого паттерна пересекается с диапазоном $F(Im_H)$, и сам отклик при этом тоже будет только низкочастотным. Если обученной НС предъявить высокочастотный паттерн Im_{Ind} , то матрица связей H_{HD} формировать никакого отклика в слое С не будет, поскольку $F(Im_{Ind})$ не пересекается с H_{HD} – сеть, обученная тому, что все люди смертны (8), не может связать это знание с Im_{Ind} .

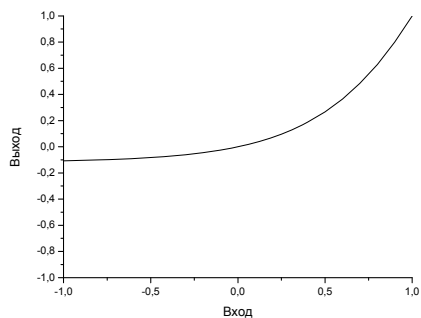


Рис.5. Активационная функция нейронов слоя С

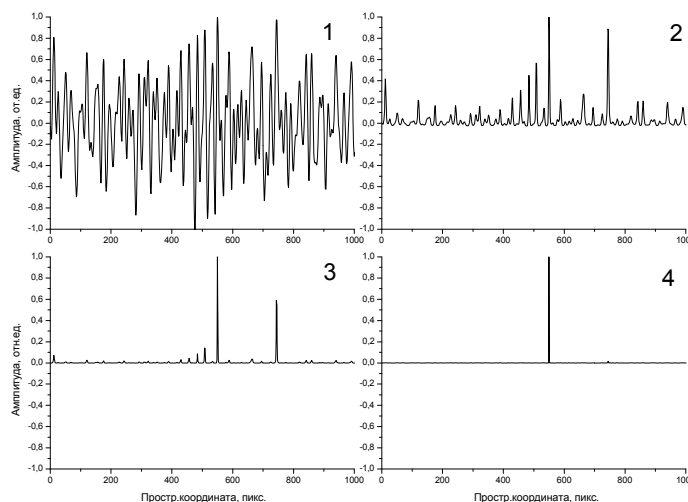


Рис.6. Исходный паттерн Im_{DInd} (1) и паттерны, полученные после 2-ой (2), 4-ой (3) и 10-ой (4) итераций

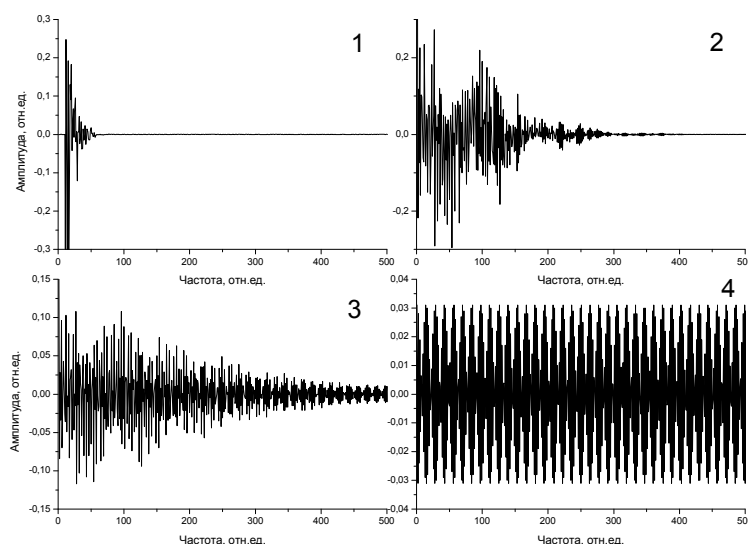


Рис.7. Спектры (Фурье-образы) 1 - исходного паттерна, 2, 3 и 4 – после 2-ой, 4-ой и 10-ой итераций, соответственно

Теперь предъявим сети результат (3), который она должна запомнить, т.е. сформировать матрицу связей между высокочастотным Im_{Ind} и высокочастотным же Im_{DInd} , представляющим свойство смертности конкретной персоны

$$H_{DInd} = F^*(Im_{Ind})F(Im_{DInd}) \quad (9)$$

Матрицы связей (голограммы) (8) и (9) разнесены по шкале частот, т.е. пространственно не пересекаются в частотной области, но физически их можно записать на одной регистрирующей среде – такая технология в голографии хорошо известна. При такой реализации матриц

связей, для реализации абдуктивного вывода необходимо преобразовать ПВР Im_{DInd} так, чтобы параллельно активировать обе матрицы связей (восстановить обе голограммы) (8) и (9), получив в слое R требуемый паттерн - решение «Данная персона – человек»

$$Im_{Person} = Im_H + Im_{Ind} \quad (10)$$

Требуемое преобразование в пространстве Фурье заключается в расширении $F(Im_{DInd})$ так, чтобы его частотный диапазон перекрывался с диапазоном H_{HD} , т.е. снова необходимо расширение спектра в область низких частот, анало-

гично тому, как это было и в случае индукции. Решение этой задачи уже описано в предыдущем разделе - за счет реализации итерационной процедуры с использованием нелинейного итерирующего отображения в слое С, позволяющего свести произвольный паттерн Im_{Dlnd} к дельта-функции. Соответственно, абдуктивный вывод реализуется в той же схеме НС, что и индуктивный – схеме двухслойной НС с двунаправленными связями между слоями, соответствующей схеме голографии Фурье с нелинейным обращением волнового фронта в корреляционной плоскости. Однако при реализации абдукции имеется и существенное отличие от индукции – голограммы не перезаписываются, а используются уже существующие. Иными словами, не происходит расширения знания теории, рассуждения возможны только в рамках уже существующей теории. Соответственно, для того, чтобы абдуктивный вывод стал возможным, теория должна быть субъективно достаточно полна – расширенный спектра, происшедшее за одну итерацию, должно быть достаточным, чтобы новый спектр пересекался, хотя бы частично, со второй голограммой.

Заключение

Таким образом, в рамках подхода к реализации образного мышления двухслойной нейронной сетью с матрицей двунаправленных связей, формируемой методом голографии Фурье, возможна реализация как достоверного (дедуктивного), так и правдоподобных выводов – индукции и абдукции. Для реализации дедуктивного вывода, в том числе, в классе нечеткозначимых немонотонных логик, достаточно однократного прохождения возбуждения через матрицу связей от слоя к слою, а активационная функция нейронов может быть линейна. Для реализации правдоподобных выводов необходимы два условия - итерационная процедура и нелинейная активационная функция нейронов в слое С. Эта итерационная процедура имеет целью в слое С сведение паттерна к дельта-функции и реализуется в рамках идеологии «победитель забирает всё».

Необходимость сведения паттерна в слое С к дельта-функции имеет также и сугубо физическую причину - невозможность формирования матрицы связей, имеющей четырехмерную область определения. Соответственно, при восста-

новлении реальной матрицы связей (голограммы) не дельта-функцией, а определенным на плоскости паттерном, в слое R будет восстанавливаться не ассоциированный паттерн, а картина, образованная суммой наложенных и смещенных друг относительно друга паттернов.

При реализации абдуктивного вывода решение задачи – новый паттерн, формируемый в слое R в результате прохождения возбуждения через матрицу связей от преобразованного в дельта-функцию паттерна в слое С. Матрица связей при этом не изменяется.

При реализации индуктивного вывода решение задачи – новая матрица связей. В результате, абдуктивный вывод расширяет знание примеров, а индуктивный – теорию, т.е. знание переднего плана. Для реализации как абдуктивного, так и индуктивного выводов необходимо сочетание определенного запаса знаний, выражающегося в достаточной заполненности спектрального диапазона паттерна внутренней репрезентации, и интеллектуальных способностей познающего субъекта в виде требуемой нелинейности активационной функции С-нейронов. Недостаток одного из компонентов, например, знаний, может быть компенсирован за счет другого – нелинейности активационной функции, достаточной для необходимого расширения спектра за одну итерацию.

Индуктивный вывод может рассматриваться также как трансформация сенсорного образа в понятие посредством разрушения внутренней коррелированности сенсорного образа (перехода к дельта-коррелированности), т.е. абстрагирования от несущественных для понятия признаков. При этом паттерн - дельта-функция в слое С может рассматриваться как понятие абсолютное, поскольку перекрывает весь возможный диапазон [сенсорный образ – вербальный образ] (белый спектр с Фурье пространстве).

Автор считает приятным долгом выразить благодарность проф. О.П. Кузнецову и проф. И.Б.Фоминых за дискуссии, способствовавшие формированию данного подхода, и критическое обсуждение статьи, а также д.т.н. А.С.Потапову за ряд критических замечаний и предложений.

Литература

1. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. М.: Физматлит, 2004.- 708с. с илл.

2. Вагин В.Н., Хотимчук К.Ю. Нахождение минимальных абдуктивных объяснений с помощью первичных импликант // Одиннадцатая Национальная конференция по искусственному интеллекту с международным участием. Труды конференции. Т.2. с.345-355.
3. Кузнецов О.П., Неклассические парадигмы в ИИ // Известия АН, Сер. Теория и системы управления, 1995, №5
4. Кузнецов О.П., Марковский А.В., Шипилина Л.Б. Голографические механизмы обработки образной информации //Ин-т проблем управления им. В.А.Трапезникова РАН. М.2007. 84 с.
5. Кузнецов О.П. Быстрые процессы мозга и обработка образов.//Новости искусственного интеллекта. 1998, №2.
6. Леутин В.П., Николаева Е.И., Функциональная асимметрия мозга. Мифы и реальность. СПб. Речь. 2005. 368С.
7. Голицын Г.А., Фоминых И.Б. Нейронные сети и экспертные системы: перспективы интеграции // Новости искусственного интеллекта. 1996. №4. с.121-145.
8. Прибрам К. Нелокальность и локализация: голографическая гипотеза о функционировании мозга в процессе восприятия и памяти // В сб. «Синергетика и психология». Вып.1. "Методологические вопросы". Изд-во МГСУ "Союз". 1997.
9. Павлов А.В. Математические модели оптических методов обработки информации // Изв. РАН. ТиСУ. 2000. №3.
10. Павлов А.В. Реализация логико-лингвистических моделей методом Фурье-голографии // Изв. РАН. ТиСУ. 2003. №2. с.118-125.
11. Павлов А.В., Интеграция логического и образного мышления методом Фурье-голографии. Десятая Национальная конференция по искусственному интеллекту с международным участием. Труды конференции. Т.2, с.644-652.
12. Алексеев А.М., Павлов А.В., Реализация нечеткой логики с исключениями методом фурье-голографии // Оптический журнал, 2008, Т. 75, №4, с.41-46.
13. Фоминых И.Б., Гибридные интеллектуальные системы на основе темпоральных логических моделей и инженерии образов. // V-я Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте» Сб. научных трудов. Т.1., с.87-96. М.: Физматлит. 2009.
14. Валькман Ю.Р., Тарасов В.Б. Круглый стол «Образ и понятие». V-я Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте». Коломна, 28-30 мая 2009.
15. Шубников Е.И., Влияние аддитивных и мультипликативных помех при корреляционном сравнении изображений // Оптика и спектроскопия, 1987, т. 62, № 3,с. 653-658
16. Павлов А.В., Влияние геометрических искажений на корреляцию изображений.// Оптика и спектроскопия, 1991, Т.70. №6. с.1337-1341.
17. Кулешов А.М., Шубников Е.И., Смаева С.А., Об оптимальности голографического согласованного фильтра.// Оптика и спектроскопия, 1986, Т.60, №6, с.1273-1276.
18. Александрина С.А., Кулешов А.М., Влияние режекции низких пространственных частот спектра на параметры сигнала в голографическом корреляторе // Оптика и спектроскопия, 1990, Т.68, №3, с.652-655.
19. Князева Е.Н., Методы нелинейной динамики в когнитивной науке // В сб. «Синергетика и психология». Вып.3. "Когнитивные процессы". Когито-Центр. 2004. с.29-48.

Павлов Александр Владимирович. Старший научный сотрудник. Санкт-Петербургского государственного университета информационных технологий, механики и оптики. Окончил Ленинградский институт точной механики и оптики в 1980 году. Кандидат технических наук, доцент. Автор 100 печатных работ. Область научных интересов: искусственный интеллект, оптика, голография, оптические информационные технологии. Pavlov@phoi.ifmo.ru.