

## ВОЗМОЖНОСТЬ РЕАЛИЗАЦИИ ПРАВДОПОДОБНЫХ РАССУЖДЕНИЙ МЕТОДОМ ГОЛОГРАФИИ ФУРЬЕ

© 2010 г. А. В. Павлов, канд. техн. наук

Санкт-Петербургский государственный университет информационных технологий, механики и оптики, Санкт-Петербург

E-mail: pavlov@phoi.ifmo.ru

В развитие ранее разработанного метода реализации нечетких логик предложен подход к реализации правдоподобных рассуждений (абдуктивного и индуктивного выводов) на двуслойной оптической нейронной сети с матрицей двунаправленных связей, реализуемой методом голографии Фурье. Определены требования к характеристикам паттернов, представляющих обрабатываемую информацию, и фазо-сопрягающего зеркала. Приведены результаты численного моделирования.

**Ключевые слова:** голография, искусственный интеллект, правдоподобные рассуждения, абдукция, индукция.

Коды OCIS: 090.0090, 100.4996, 070.4560

Поступила в редакцию 16.11.2009

### Введение

Хорошо известно наличие ряда глубоких аналогий между принципами функционирования мозга с одной стороны, и физическими явлениями и механизмами, лежащими в основе голографии, с другой [1–4]. Эти аналогии послужили основанием формирования голографической парадигмы (ГП) в когнитивных науках и искусственном интеллекте (ИИ). Хорошо известны реализации голографической ассоциативной памяти [5]. В работах [6–10] разработана методика реализации методом голографии Фурье логического вывода в классе немонотонных нечетко-значимых логик.

Вместе с тем, основной функцией интеллекта как биологического, так и искусственного, является рассуждение, под которым принято понимать построение последовательности аргументов, вынуждающих принятие некоторого утверждения [11]. Принято выделять два типа рассуждений или выводов – достоверные, к которым относится дедуктивный вывод, и правдоподобные, к которым относятся абдукция и индукция [11]. Логика имеет дело по преимуществу с дедуктивным выводом, который привлекал внимание исследователей на протяжении веков,

начиная, как минимум, с Аристотеля. Правдоподобные рассуждения, включая их реализацию методами ИИ, до последнего времени привлекали существенно меньшее внимание исследователей и разработчиков, чем достоверные рассуждения [11, 12]. При этом многие исследователи отмечали, что правдоподобные рассуждения являются не просто атрибутом человеческого интеллекта, но и важнейшей, неотъемлемой частью процесса познания. Поэтому представляет безусловный интерес возможность их реализации в рамках концепции образного мышления методом голографии.

### Биологически мотивированный подход и нейросетевая модель

#### 1. Подход к задаче

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 [1]. Эти образы суть не философские абстракции, но физически реальные картины нейронной активности коры головного мозга –

паттерны внутренней репрезентации (ПВР) воспринимаемой или вспоминаемой информации. Поскольку ПВР как картина нейронной активности представляет собой пространственное распределение электрических потенциалов по коре головного мозга [13], то примем в качестве аналога биологического ПВР (применительно к рассматриваемому методу) обрабатываемое голографической схемой изображение  $Im(x, y)$  как определенное на плоскости распределение амплитуд (для поля) или пропусканий (для транспаранта), не детализируя пока его характеристики. В дальнейшем, для простоты выражений, обозначения координат опустим.

Для большей наглядности используем в дальнейших построениях классический пример классификации рассуждений (по Ч. Пирсу) [11]:

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

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

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, так же 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 имени персоны.

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 [1–3, 14–15]. 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, т. е. продукт деятельности мозга. Образ всегда характеризуется определенной целостностью, это гештальт [2, 3, 15]. Понятие же формируется в процессе разрушения гештальта, вычленения из него только существенных для понятия признаков и абстрагирования от несущественных, но неотъемлемых 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.

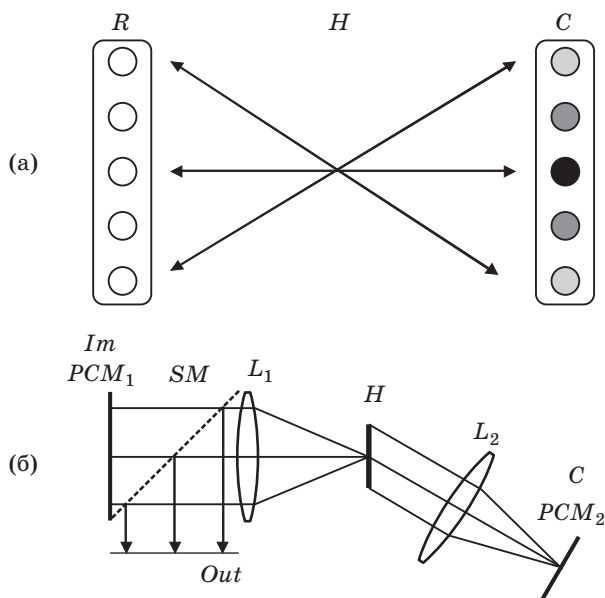
Таким образом, правило (1) связывает два понятия –  $Im_H$  и  $Im_D$ .

Случай (2) связывает ПВР – образ, представляющий только сугубо индивидуальные свойства Сократа, который обозначим  $Im_{IndСократ}$ , с понятием  $Im_H$ . Здесь важно, что по условиям примера ПВР  $Im_{IndСократ}$  (сугубо сенсорный образ) не включает в себя ПВР  $Im_H$  (понятие), поскольку о том, что Сократ – человек, мы узнаем только в результате связи двух этих ПВР. В рамках приведенного примера такая схема выглядит несколько надуманной, но для других, более сложных примеров, относящихся к методологии научного познания [11], она вполне реальна. Кроме того, этот простой пример нужен для понимания подхода к решению поставленной задачи.

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

Рассмотрим возможность реализации этих связок двуслойной НС с двунаправленными связями, приведенной на рис. 1. Выбор этой простой архитектуры обусловлен тем, что она адекватна реальной схеме голографии Фурье [6–10].

В таблице приведен вариант представления ПВР слоями НС, позволяющий реализовать как



**Рис. 1.** Модель двухслойной нейронной сети (а) с двунаправленными связями ( $R$  и  $C$  – нейронные слои,  $H$  – матрица связей нейронных слоев), адекватная схеме голографии Фурье (б) с обращением волновых фронтов в плоскостях изображений и корреляций ( $Im$  – плоскость изображений,  $C$  – плоскость корреляций,  $L_1$ ,  $L_2$  – фурье-преобразующие линзы,  $H$  – голограмма Фурье,  $SM$  – полупрозрачное зеркало, формирующее выходную плоскость  $Out$ ,  $PCM_1$  и  $PCM_2$  – фазо-сопрягающие зеркала, помещенные в плоскости изображений и корреляций соответственно).

Представление паттернов внутренней репрезентации в слоях искусственной нейронной сети, реализуемой методом голографии Фурье

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

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

## 2. Выбор метода представления информации на примере дедуктивного вывода

### 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  $Im_{Ind\text{Сократ}}$  реализовать цепочку ассоциаций

$$Im_{Ind\text{Сократ}} \rightarrow Im_H \rightarrow Im_D,$$

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  $Im_{Ind\text{Сократ}} \rightarrow Im_H$  в предыдущих рассуждениях был отвергнут, то в рамках принятого подхода (таблица, 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, мужчина 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 в рамках нашей модели НС и ее оптической реализации используем шкалу частот в пространстве Фурье, в котором формируется матрица связей нейронных слоев  $H$  в виде голограммы Фурье (плоскость  $H$  на рис. 1б). В этом случае 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

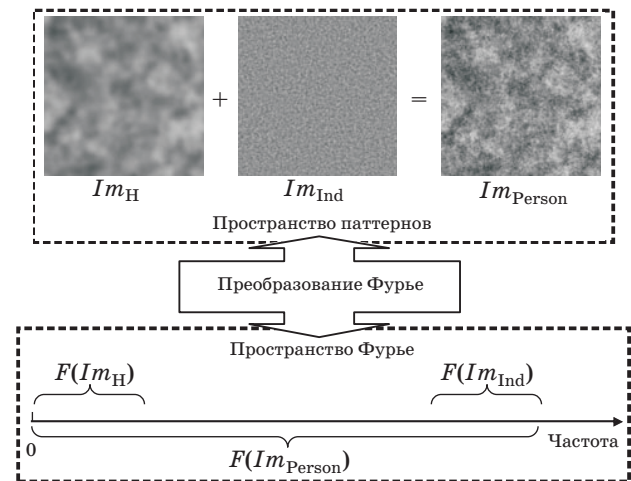
формирует оптическая система (глаз, например), но результат преобразования этого изображения сенсором и сенсорным трактом. ПВР как картина распределения нейронной активности совершенно не похож на изображение. В этой связи отметим, что выбор шкалы частот в качестве шкалы общности свойств имеет своим обоснованием не столько вышеприведенную интуитивно понятную интерпретацию из области зрительного восприятия, сколько то, что ПВР, представляющие более 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 корреляционной функции к искажениям объектного изображения (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,  $r_x, r_y$  – радиусы корреляции [16, 17]. А больший радиус корреляции определяется именно большим удельным весом низких частот в спектре [18, 19].

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 ее 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 общности свойств представлены субпаттернами  $Im_H$  и  $Im_{Ind}$ , а между ними на шкале общности представлены другие субпаттерны, представляющие иные свойства и признаки персоны (пол, возраст, etc.). Эти субпаттерны накладываются друг на друга в пространстве паттернов (слое  $R$  на рис. 1а, соответствующем плоскости изображений  $Im$  на рис. 1б), но их образы Фурье разнесе-



**Рис. 2.** Схема представления случая (2) “Данная персона – человек”. Субпаттерн  $Im_{Ind}$  представляет индивидуальные свойства персоны, а субпаттерн  $Im_H$  – свойство “человек”.

ны в пространстве Фурье (плоскость голограммы  $H$  на рис. 1б), поскольку занимают разные частотные диапазоны (рис. 2)

$$Im_{Person} = Im_H + \dots + Im_{Ind}, \quad (4)$$

$$F(Im_{Person}) = F(Im_H) \cup \dots \cup F(Im_{Ind}).$$

## 2.2. Характеристики паттерна “смертен” при реализации дедукции

Обучение сети (рис. 1) правилу (1) заключается в формировании матрицы связей  $H$ , связывающей низкочастотный субпаттерн  $Im_H$  с паттерном, представляющим в слое  $C$  свойство “смертен”  $Im_D$

$$H_{HD} = F^*(Im_H)F(Im_D), \quad (5)$$

где  $F$  обозначает преобразование Фурье, а астериск – комплексное сопряжение. Выражение (5) соответствует правилу обучения Хэбба и описывает голограмму Фурье, реализующую матрицу связей (5).

В пространстве Фурье частотный диапазон паттерна  $Im_D$  должен пересекаться с частотным диапазоном субпаттерна  $Im_H$  – в противном случае голограмма (5) сформирована быть не может. Поскольку выше было определено, 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 не влияет.

Случай (2) реализуется предъявлением обученной сети (восстановлением голограммы) паттерна персоны (4), включающего в себя субпаттерн  $Im_H$ , в результате сеть формирует отклик

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

где символы  $*$  и  $\otimes$  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  $F(Im_{Ind})F^*(Im_H)$  является помехой, а корреляционный алгоритм, как известно, осуществляет оптимальное по критерию отношения сигнал/помеха выделение сигнала из шума, т. е. также получаем требуемый результат “данная персона смертна”, но с меньшим отношением сигнал/помеха [16]

$$Out = F[F(Im_{Person})F^*(Im_H)F(Im_D)] = Im_D.$$

### 3. Реализация индуктивного вывода

При реализации индуктивного вывода сеть должна обучаться примерами результатов (3), связывающих паттерн персоны (4) с ПВР свойства “смертен”, которое рассматривается здесь (по условиям задачи индукции) как сугубо индивидуальное, не обладающее общностью – мы знаем из единичного опыта, что данная персона смертна, но еще не знаем, что свойство смертности всеобщее. Обозначим образ, представляющий такое, сугубо индивидуальное, свойство добавлением индекса Ind, т. е.  $Im_{DInd}$ . Формируемая при обучении результатом (3) голограмма описывается выражением

$$\begin{aligned} H_{DInd} &= F^*(Im_{Person})F(Im_{DInd}) = \\ &= F^*(Im_H)F(Im_{DInd}) + \\ &+ F^*(Im_{Ind})F(Im_{DInd}) = F^*(Im_{Ind})F(Im_{DInd}). \end{aligned} \quad (6)$$

Если при реализации дедукции частотный диапазон  $Im_D$  пересекался с диапазоном  $Im_H$ , т. е.  $Im_D$  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. Задача 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 накладывается. Частный случай решения этой задачи в пространстве паттернов – преобразование произвольного паттерна в  $\delta$ -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 итерированного 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.

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 рисунка 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. В терминах голографии, снова, как и при реализации дедукции, получается схема голографии Фурье, но с нелинейным обращением волнового фронта в плоскости корреляций.

Если активационная функция С-нейронов (экспозиционная характеристика фазо-сопрягающего зеркала в корреляционной плоскости рис. 1) нелинейна, 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  $\delta$ -функции

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

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

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

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

Большинство реальных голографических регистрирующих сред имеют участок экспозиционной характеристики, удовлетворяющей этому условию; как правило, это область недоэкспозиций и малых экспозиций [18, 19].

Подчеркнем важность неединственности результата (3), которым обучается НС. В принципе, индуктивно обобщение возможно и на единственном примере. Однако здесь не накладывалось никаких ограничений на  $Im_{DInd}$  и, если  $Im_{DInd}$  окажется гармоническим колебанием, то итерационная процедура не позволит свести его к  $\delta$ -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 гармоникой, но другой частоты, ведет к появлению комбинационных, в том числе разностных, частот.

Способность к обобщению зависит как от активационной функции  $C$ -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. Если свойство смертности неактуально для познающего субъекта, то и “яркость”  $Im_{DInd}$  не велика по сравнению с другими образами, одновременно с  $Im_{DInd}$ , представленными в слое  $C$ . 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 соответствует вышеупомянутому [1–3] отличию образа от понятия в плане наличия у первого целостности,

т. е. внутренней коррелированности. Переход от произвольного паттерна  $Im_{DInd}$  к  $\delta$ -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  $\delta$ -функции, а переход в пространстве Фурье к постоянной спектральной плотности. Именно постоянная спектральная плотность обеспечивает общность понятия, его применимость к любому субпаттерну из совокупности (4), составляющей паттерн персоны.

Поскольку здесь введено новое (по сравнению с  $Im_D$ ) понятие смертности вообще, характеризующееся  $\delta$ -коррелированностью, а введенное ранее понятие смертности  $Im_D$  обладало не нулевым, а большим радиусом корреляции, то обозначим новое понятие паттерном  $Im_{DExt}$ . Понятие  $Im_{DExt}$  оказывается действительно общим, поскольку применимо как к сенсорным образам  $Im_{Ind}$ , так и к вербальным  $Im_H$ . Постоянная спектральная плотность, в свою очередь, через теорему Винера–Хинчина обуславливает  $\delta$ -коррелированность вне зависимости от конкретного вида паттерна в слое  $C$ : это может быть как единичная  $\delta$ -функция, так и  $\delta$ -коррелированный паттерн.

Вместе с тем, переход в слое  $C$  от произвольного паттерна  $Im_{DInd}$  именно к единичной  $\delta$ -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-мерную область определения, что физически нереализуемо. В случае же единственной  $\delta$ -функции в слое  $C$ , матрица связей определена на плоскости, т. е. физически реализуема в виде тонкой голограммы Фурье.

Таким образом, индуктивный вывод реализуется в рассмотренной архитектуре НС рисунка 1. Если для реализации дедукции было достаточно однократного прохождения света от слоя  $R$  к слою  $C$  и, соответственно, однонаправленных связей, то для реализации индукции необходимы двунаправленные связи, позволяющие реализовать итерационную процедуру. Эта двунаправленность связей реализуется помещением в корреляционную плоскость и плоскость изображений фазо-сопрягающих зеркал.

На рис. 3 приведен пример экспозиционной характеристики фазо-сопрягающего зеркала в корреляционной плоскости (нелинейной акти-

вационной функции  $C$ -нейронов), на рис. 4 – результаты трансформации паттерна с ограниченным экспоненциальным спектром к  $\delta$ -функции, а на рис. 5 – соответствующие им действительные части фурье-образов.

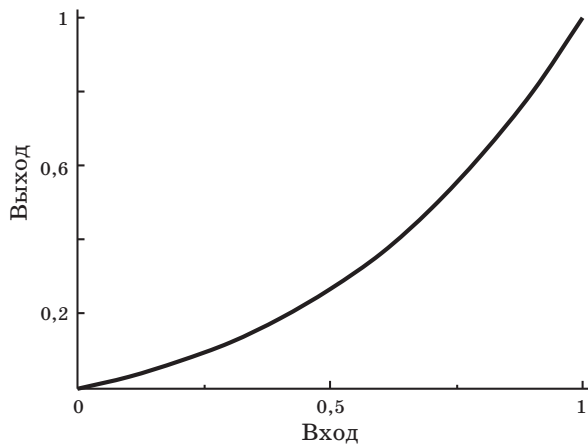


Рис. 3. Активационная функция нейронов слоя  $C$  (экспозиционная характеристика фазо-сопрягающего зеркала в корреляционной плоскости).

#### 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 рассматриваемого подхода – задача формирования в составе  $Im_{Person}$  низкочастотного субпаттерна  $Im_H$ .

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

$$H_{HD} = F^*(Im_H)F(Im_D). \quad (7)$$

Напомним, что (7) – низкочастотная голограмма, поскольку  $F(Im_H)$  локализован только в области низких частот (рис. 2). При предъявлении обученной таким образом сети какого-либо паттерна в одном из слоев ( $R$  или  $C$ ) отклик в другом слое ( $C$  или  $R$  соответственно) будет сфор-

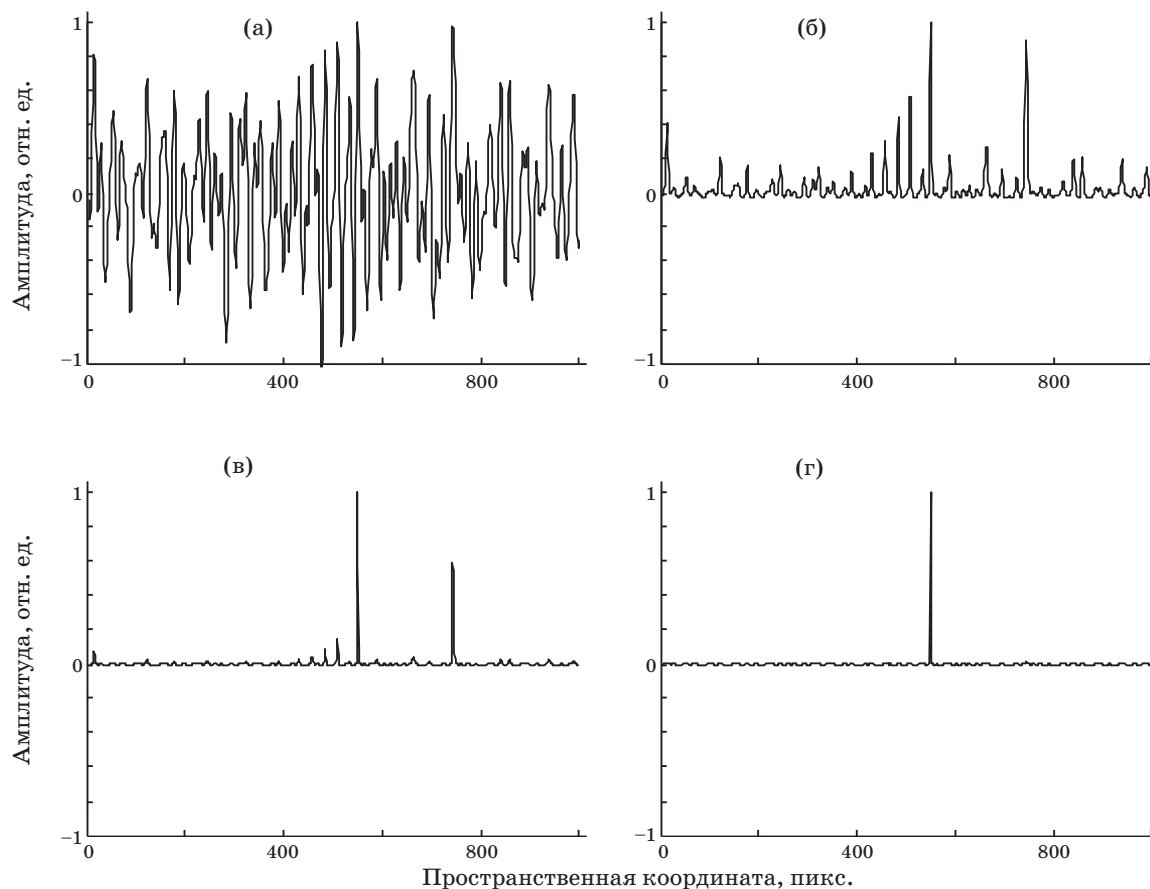
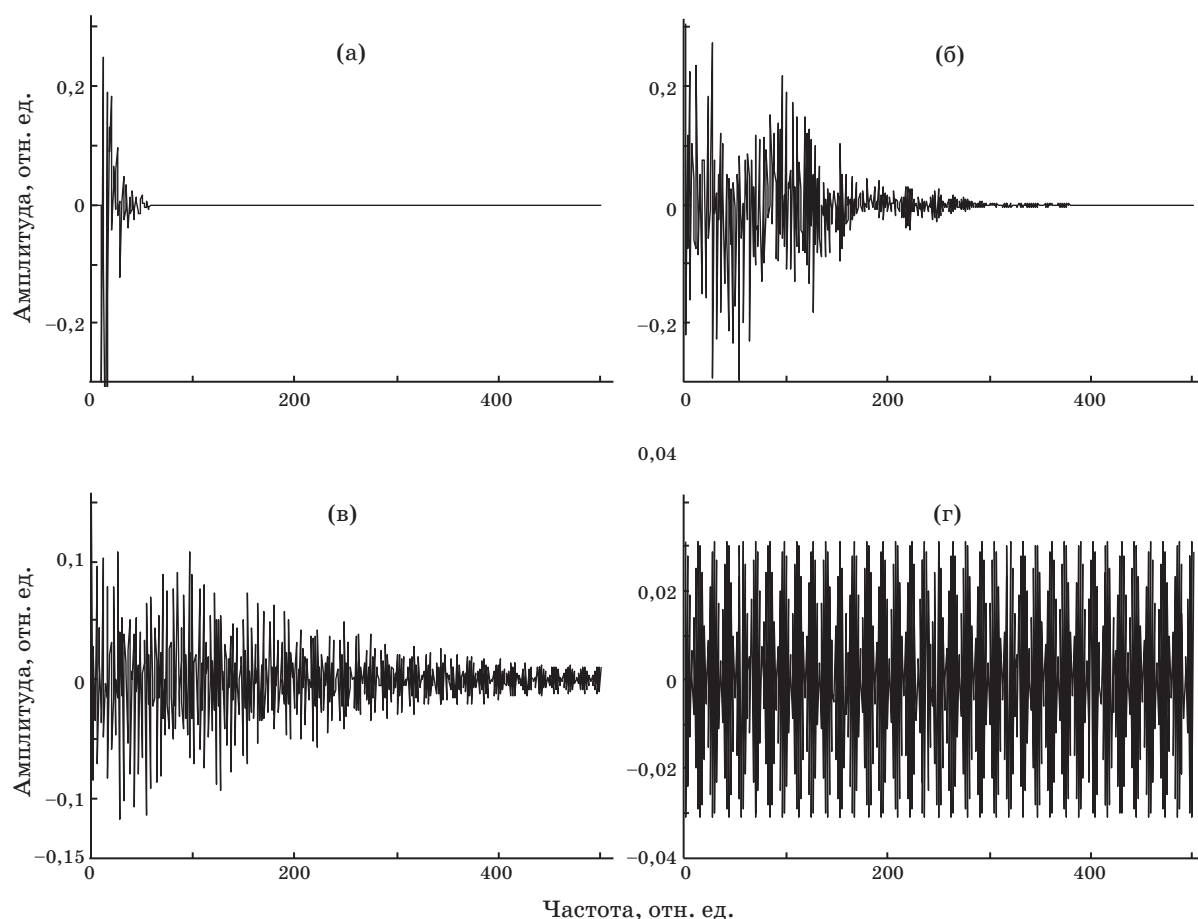


Рис. 4. Исходный паттерн  $Im_{DInd}$  (а) и паттерны, полученные после 2-й (б), 4-й (в) и 10-й (г) итераций.



**Рис. 5.** Спектры (действительные части фурье-образов). Исходный паттерн (а) и после 2-й (б), 4-й (в) и 10-й (г) итераций соответственно.

мирован 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 голограмма  $H_{HD}$  формировать никакого отклика в слое  $C$  не будет, поскольку  $F(Im_{Ind})$  не пересекается с  $H_{HD}$ : сеть, обученная тому, что все люди смертны (7), не может связать это знание с  $Im_{Ind}$ .

Теперь предъявим сети результат (3), который она запомнит, т. е. запишем голограмму, связывающую высокочастотный  $Im_{Ind}$  и высокочастотный же  $Im_{DInd}$ , представляющий свойство смертности конкретной персоны

$$H_{DInd} = F^*(Im_{Ind})F(Im_{DInd}). \quad (8)$$

Голограммы (7) и (8) разнесены по шкале частот, т. е. пространственно не пересекаются в частотной области, но физически их можно записать на одной регистрирующей среде. Тогда, для реализации абдукции необходимо преобразовать ПВР  $Im_{DInd}$  так, чтобы параллельно вос-

становить им обе голограммы (7) и (8), получив в слое  $R$  требуемый паттерн – решение “Данная персона – человек”

$$Im_{Person} = Im_H + Im_{Ind}. \quad (9)$$

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 аналогично 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  $C$ , позволяющего свести произвольный паттерн  $Im_{DInd}$  к  $\delta$ -функции. Соответственно, абдуктивный вывод реализуется в той же схеме НС, что и индуктивный – схеме двуслойной НС с двунаправленными связями между слоями, соответствующей схеме голографии Фурье с нелинейным обращением волнового фронта в корреляционной плоскости.



## Заключение

Таким образом, в рамках подхода к реализации образного мышления двуслойной 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 достаточно однократного прохождения возбуждения через матрицу связей от слоя к слою, а активационная функция нейронов (характеристики сенсоров) может быть линейна. Для реализации правдоподобных выводов необходимы два условия – итерационная процедура и нелинейная экспозиционная характеристика фазо-сопрягающего зеркала (нелинейная активационная функция нейронов) в слое *C*. Эта итерационная процедура, реализующая идеологию “победитель забирает все”, имеет целью сведение паттерна в слое *C* к  $\delta$ -функции.

Необходимость сведения паттерна в слое *C* к  $\delta$ -функции имеет также и сугубо физическую причину – невозможность формирования голограммы как матрицы связей, имеющей четырехмерную область определения. Соответственно, при восстановлении реальной голограммы не  $\delta$ -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 – это паттерн, формируемый в слое *R* в результате прохождения возбуждения через матрицу связей от преобразованного в  $\delta$ -функцию паттерна в слое *C*. При реализации индуктивного вывода решение задачи – новая голограмма (матрица связей). В результате, абдуктивный вывод расширяет знание примеров, а индуктивный – теорию, т. е. знание переднего плана.

Индуктивный вывод может рассматриваться также как трансформация сенсорного образа в понятие посредством разрушения внутренней коррелированности сенсорного образа (перехода к  $\delta$ -коррелированности), т. е. абстрагирования от несущественных для понятия признаков. При этом паттерн –  $\delta$ -функция в слое *C*, может рассматриваться как понятие абсолютное, поскольку перекрывает весь возможный диапазон смыслов (белый спектр в пространстве Фурье).

Автор считает приятным долгом выразить благодарность профессору О.П. Кузнецову и

профессору И.Б. Фоминых за дискуссии, способствовавшие формированию данного подхода, и критическое обсуждение статьи.

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

## ЛИТЕРАТУРА

1. Кузнецов О.П., Марковский А.В., Шипилина Л.Б. Голографические механизмы обработки образной информации. М.: Ин-т проблем управления им. В.А. Трапезникова РАН, 2007. 84 с.
2. Кузнецов О.П. Голографические модели обработки информации в нейронных сетях // Докл. АН. 1992. Т. 324. № 3. С. 537–540.
3. Голицын Г.А., Фоминых И.Б. Нейронные сети и экспертные системы: перспективы интеграции // Новости искусственного интеллекта. 1996. № 4. С. 121–145.
4. Прибрам К. Нелокальность и локализация: голографическая гипотеза о функционировании мозга в процессе восприятия и памяти // “Синергетика и психология”. В. 1. “Методологические вопросы”. М.: МГСУ “Союз”, 1997. С. 136.
5. Owechko Y. Nonlinear holographic associative memories // IEEE Journ. Quant. Electron. 1989. V. 25. № 3. P. 619–634.
6. Павлов А.В. Применение методов Фурье-голографии для построения логических процессоров // Оптический журнал. 2002. Т. 69. № 10. С. 42–48.
7. Павлов А.В. Реализация логико-лингвистических моделей методом Фурье-голографии // Изв. РАН. Серия: Теория и системы управления. 2003. № 2. С. 118–125.
8. Алексеев А.М., Константинов А.М., Павлов А.В. Использование метода Фурье-голографии для моделирования принципа образности мышления // Оптический журнал. 2006. Т. 73. № 9. С. 77–82.
9. Павлов А.В. Интеграция логического и образного мышления методом Фурье-голографии // X Нац. конф. по искусственному интеллекту с международным участием. Обнинск, 25–28.09.2006. Труды конф. М.: Физматлит, 2006. Т. 2. С. 644–652.
10. Алексеев А.М., Павлов А.В. Реализация нечеткой логики с исключениями методом Фурье-голографии // Оптический журнал. 2008. Т. 75. № 4. С. 41–46.
11. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. М.: Физматлит, 2004. 708 с.
12. Вагин В.Н., Хотимчук К.Ю. Нахождение минимальных абдуктивных объяснений с помощью первичных импликант // XI Нац. конф. по искусственному интеллекту с международным участием. Дубна, 29.09–03.10.2008. Труды конф. М.: Физматлит, 2008. Т. 2. С. 345–355.

13. *Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р.* Модели динамики нейронной активности при обработке информации мозгом – итоги “десятилетия” // УФН. 2002. Т. 172. № 10. С. 1189–1214.
  14. *Фоминых И.Б.* Гибридные интеллектуальные системы на основе темпоральных логических моделей и инженерии образов // V Межд. научно-практ. конф. “Интегрированные модели и мягкие вычисления в искусственном интеллекте”. Сб. научных трудов. М.: Физматлит, 2009. Т. 1. С. 87–96.
  15. *Валькман Ю.Р., Тарасов В.Б.* Устное сообщение на круглом столе “Образ и понятие” // V Межд. научно-практ. конф. “Интегрированные модели и мягкие вычисления в искусственном интеллекте”. Коломна, 28–30 мая 2009.
  16. *Шубников Е.И.* Влияние аддитивных и мультипликативных помех при корреляционном сравнении изображений // Опт. и спектр. 1987. Т. 62. № 3. С. 653–658.
  17. *Павлов А.В.* Влияние геометрических искажений на корреляцию изображений // Опт. и спектр. 1991. Т. 70. № 6. С. 1337–1341.
  18. *Кулешов А.М., Шубников Е.И., Смаева С.А.* Об оптимальности голографического согласованного фильтра // Опт. и спектр. 1986. Т. 60. № 6. С. 1273–1276.
  19. *Александрина С.А., Кулешов А.М.* Влияние режекции низких пространственных частот спектра на параметры сигнала в голографическом корреляторе // Опт. и спектр. 1990. Т. 68. № 3. С. 652–655.
-