

УДК 303.732.4

UDC 303.732.4

**СИСТЕМНО-КОГНИТИВНЫЙ АНАЛИЗ
ИЗОБРАЖЕНИЙ (обобщение, абстрагирование,
классификация и идентификация)****SYSTEMIC-COGNITIVE ANALYSIS OF
IMAGES (generalization, abstraction, classification
and identification)**

Луценко Евгений Вениаминович
д.э.н., к.т.н., профессор
Кубанский государственный аграрный универси-
тет, Краснодар, Россия

Lutsenko Evgeny Veniaminovich
Dr. Sci.Econ., Cand. Tech.Sci., professor
Kuban State Agrarian University, Krasnodar, Russia

В статье рассматривается применение системно-когнитивного анализа, его математической модели – системной теории информации и программного инструментария – системы «Эйдос» для синтеза обобщенных изображений классов, их абстрагирования, классификации обобщенных изображений (кластеры и конструкторы) сравнения конкретных изображений с обобщенными образами (идентификация)

In the article application of systemic-cognitive analysis, its mathematical model - the system theory of the information and program toolkit - systems "Eidoses" for synthesis of the generalized images of classes, their abstraction, classification of the generalized images (clusters and constructs) comparisons of concrete images with the generalized images (identification) are examined

Ключевые слова: СИСТЕМНО-КОГНИТИВНОГО АНАЛИЗА, ЕГО МАТЕМАТИЧЕСКОЙ МОДЕЛИ – СИСТЕМНОЙ ТЕОРИИ ИНФОРМАЦИИ И ПРОГРАММНОГО ИНСТРУМЕНТАРИЯ – СИСТЕМЫ «ЭЙДОС» ДЛЯ СИНТЕЗА ОБОБЩЕННЫХ ИЗОБРАЖЕНИЙ, ИХ АБСТРАГИРОВАНИЯ, КЛАССИФИКАЦИИ ОБОБЩЕННЫХ ИЗОБРАЖЕНИЙ (КЛАСТЕРЫ И КОНСТРУКТЫ) СРАВНЕНИЯ КОНКРЕТНЫХ ИЗОБРАЖЕНИЙ С ОБОБЩЕННЫМИ (ИДЕНТИФИКАЦИЯ).

Keywords: SYSTEMIC-COGNITIVE ANALYSIS, ITS MATHEMATICAL MODEL - THE SYSTEM THEORY OF THE INFORMATION AND PROGRAM TOOLKIT - SYSTEMS "EIDOSES" FOR SYNTHESIS OF THE GENERALIZED IMAGES OF CLASSES, THEIR ABSTRACTION, CLASSIFICATION OF THE GENERALIZED IMAGES (CLUSTERS AND CONSTRUCTS) COMPARISONS OF CONCRETE IMAGES WITH THE GENERALIZED IMAGES (IDENTIFICATION)

Несомненный научный и практический интерес представляет *синтез обобщенных изображений* на основе ряда конкретных (задача-1). При этом в результате обобщения выясняется *ценность* признаков изображений для их дифференциации, а также *степень характерности* тех или иных признаков для конкретных изображений. Это позволяет без ущерба для адекватности модели *удалить* из нее малоценные признаки, т.е. осуществить *синтез абстрактных изображений* (задача-2), что обеспечивает в последующем сокращение затрат различных видов ресурсов на сбор и обработку графической информации. Над обобщенными изображениями возможны операции классификации, объединения наиболее сходных из них в кластеры и формирования систем наиболее различных кластеров, т.е. конструкторов (задача-3). Можно также количественно сравнивать конкретные изображения с обобщенными, т.е. идентифицировать эти конкретные изображения (задача-4).

Рассмотрим решение этих задач в системно-когнитивном анализе (СК-анализ) [1, 2, 3, 4, 5]. При этом в качестве примера для решения пере-

численных задач выберем изображения десятичных цифр, используемые для написания почтовых индексов (рисунок 1):

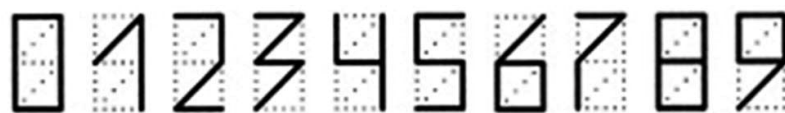


Рисунок 1. Образец написания цифр почтового индекса (ГОСТ Р 51506-99)¹.

Выбор данного примера связан прежде всего с его простотой, но тем ни менее он позволяет рассмотреть все методы и инструменты СК-анализа, которые имеют общее значение и могут быть применены для решения поставленных задач при обработке более сложных изображений.

Для решения 1-й задачи прежде всего необходимо разбить исходные изображения на элементы, из которых они составлены и закодировать как сами эти элементы, так и изображения с использованием справочников элементов.

Рассмотрение самих способов выделения и кодирования элементов изображений не входит в задачи данной статьи. Отметим лишь, что на наш взгляд эти методы могут быть основаны на метрических и топологических свойствах изображений и оба этих подхода могут базироваться на выделении элементов на основе высоких градиентов (скорости *изменения*, т.е. 1-й производной) цвета, его насыщенности, яркости и кривизны распределения этих параметров в пространстве.

В ряде конкретных случаев эта работа уже проведена. Например, в МВД существуют, специальные атласы для разработки фотороботов лиц, содержащие изображения таких элементов лиц как носы, губы, брови, глаза, лбы, морщины, залысины, прически, усы, подбородки, бороды, уши и т.д. и т.п.

Очевидно, что рассматриваемый нами шрифт специально сконструирован специалистами таким образом, чтобы было вполне очевидно из каких элементов состоят цифры, т.к. они выделены яркостью и высокой кривизной на границах элементов (четко выраженные «углы»). Для нас существенно, какие именно коды будут иметь эти элементы, поэтому пронумеруем их так, как проще (рисунок 2).

¹ <http://www.prodtp.ru/index.php?act=recipes&CODE=03&id=51>

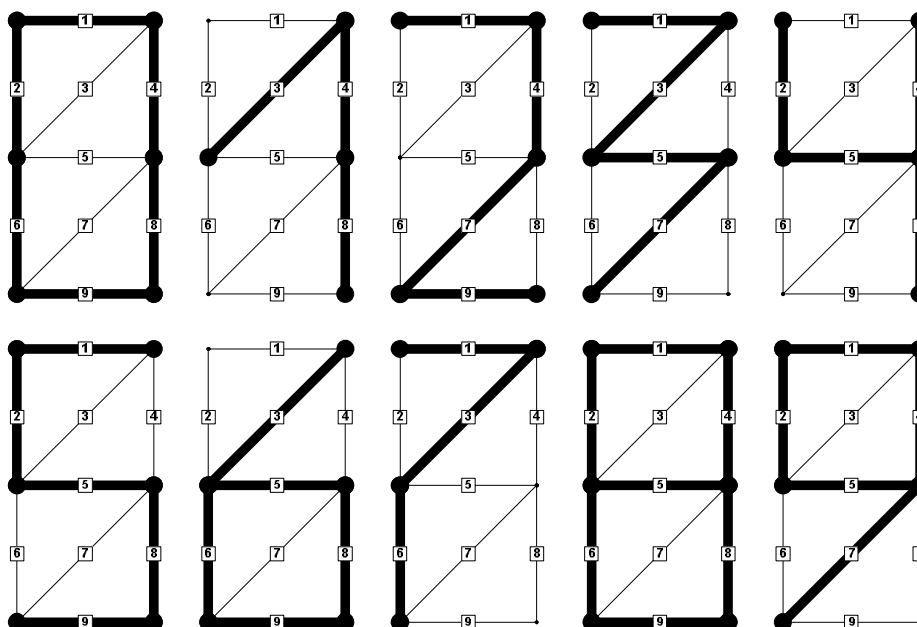


Рисунок 2. Принцип кодирования изображений десятичных цифр, состоящих из пронумерованных элементов

С использованием системы кодирования элементов цифр, представленной на рисунке 2, сформируем следующий справочник элементов (таблица 1):

Таблица 1 – СПРАВОЧНИК ЭЛЕМЕНТОВ (ПРИЗНАКОВ) ЦИФР

KOD	NAME
1	Элемент 1
2	Элемент 2
3	Элемент 3
4	Элемент 4
5	Элемент 5
6	Элемент 6
7	Элемент 7
8	Элемент 8
9	Элемент 9

Сами цифры кодируются в виде последовательности кодов элементов, из которых они состоят, при этом порядок кодов несущественен, т.к. код каждого элемента однозначно определяет его вид и положение в матрице символа (таблица 2).

Таблица 2 – КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ ДЕСЯТИЧНЫХ ЦИФР

Цифра	Признаки (элементы) цифр						
0	1	2	4	6	8	9	
1		3	4	8			
2	1		4	7	9		
3	1		3	5	7		
4	2		4	5	8		
5	1	2		5	8	9	
6	3		5	6	8	9	
7	1		3	6			
8	1	2	4	5	6	8	9
9	1	2	4	5	7		

В качестве справочника классов, формируемых на основе примеров конкретных изображений цифр, закодированных в таблице 1, примем как сами цифры, так и два обобщенных класса: «Четные» и «Не четные» (таблица 3):

Таблица 3 – СПРАВОЧНИК КЛАССОВ

KOD	NAME
1	Цифра 0
2	Цифра 1
3	Цифра 2
4	Цифра 3
5	Цифра 4
6	Цифра 5
7	Цифра 6
8	Цифра 7
9	Цифра 8
10	Цифра 9
11	Четная цифра
12	Не четная цифра

С использованием справочника классов (таблица 3) и закодированных изображений цифр (таблица 2) формируется обучающая выборка (таблица 4):

Таблица 4 – ОБУЧАЮЩАЯ ВЫБОРКА

Код	Наименование объекта	Коды классов		Коды признаков							
		1	11	1	2	4	6	8	9		
1	Цифра 0	1	11	1	2	4	6	8	9		
2	Цифра 1	2	12	3	4	8					
3	Цифра 2	3	11	1	4	7	9				
4	Цифра 3	4	12	1	3	5	7				
5	Цифра 4	5	11	2	4	5	8				
6	Цифра 5	6	12	1	2	5	8	9			
7	Цифра 6	7	11	3	5	6	8	9			
8	Цифра 7	8	12	1	3	6					
9	Цифра 8	9	11	1	2	4	5	6	8	9	
10	Цифра 9	10	12	1	2	4	5	7			

В систему «Эйдос», являющуюся инструментарием СК-анализа, вводятся справочник классов, признаков и обучающая выборка (таблицы 3, 1 и 4), а затем осуществляется синтез семантической информационной модели (СИМ-2) [1, 2, 3, 4, 5].

В результате формируются матрица абсолютных и относительных частот, а также матрица знаний, содержащая информацию о том, какая цифра предъявлена, если установлено, что в нее входит *i*-й элемент (таблицы 5, 6, 7):

Таблица 5 – МАТРИЦА АБСОЛЮТНЫХ ЧАСТОТ

Код элемента	Наименование элемента	Наименования и коды классов												Всего:
		Цифра 0	Цифра 1	Цифра 2	Цифра 3	Цифра 4	Цифра 5	Цифра 6	Цифра 7	Цифра 8	Цифра 9	Четные	Не четные	
		1	2	3	4	5	6	7	8	9	10	11	12	
1	Элемент 1	1	0	1	1	0	1	0	1	1	1	3	4	14
2	Элемент 2	1	0	0	0	1	1	0	0	1	1	3	2	10
3	Элемент 3	0	1	0	1	0	0	1	1	0	0	1	3	8
4	Элемент 4	1	1	1	0	1	0	0	0	1	1	4	2	12
5	Элемент 5	0	0	0	1	1	1	1	0	1	1	3	3	12
6	Элемент 6	1	0	0	0	0	0	1	1	1	0	3	1	8
7	Элемент 7	0	0	1	1	0	0	0	0	0	1	1	2	6
8	Элемент 8	1	1	0	0	1	1	1	0	1	0	4	2	12
9	Элемент 9	1	0	1	0	0	1	1	0	1	0	4	1	10
	Всего:	1	1	1	1	1	1	1	1	1	1	5	5	20

Таблица 6 – МАТРИЦА УСЛОВНЫХ ВЕРОЯТНОСТЕЙ (% , СИМ-2)

Код эл.	Наименование элемента	Наименования и коды классов												Безусловная вероятность
		Цифра 0	Цифра 1	Цифра 2	Цифра 3	Цифра 4	Цифра 5	Цифра 6	Цифра 7	Цифра 8	Цифра 9	Четные	Не четные	
		1	2	3	4	5	6	7	8	9	10	11	12	
1	Элемент 1	100	0	100	100	0	100	0	100	100	100	60	80	70
2	Элемент 2	100	0	0	0	100	100	0	0	100	100	60	40	50
3	Элемент 3	0	100	0	100	0	0	100	0	0	0	20	60	40
4	Элемент 4	100	100	100	0	100	0	0	0	100	100	80	40	60
5	Элемент 5	0	0	0	100	100	100	100	0	100	100	60	60	60
6	Элемент 6	100	0	0	0	0	0	100	100	100	0	60	20	40
7	Элемент 7	0	0	100	100	0	0	0	0	0	100	20	40	30
8	Элемент 8	100	100	0	0	100	100	100	0	100	0	80	40	60
9	Элемент 9	100	0	100	0	0	100	100	0	100	0	80	20	50
	Всего:	1	1	1	1	1	1	1	1	1	1	5	5	20

Таблица 7 – МАТРИЦА ЗНАНИЙ (Бит×100, СИМ-2)

Код эл.	Наименование элемента	Наименования и коды классов												Ср.кв.откл.
		Цифра 0	Цифра 1	Цифра 2	Цифра 3	Цифра 4	Цифра 5	Цифра 6	Цифра 7	Цифра 8	Цифра 9	Четные	Не четные	
		1	2	3	4	5	6	7	8	9	10	11	12	
1	Элемент 1	52		52	52		52		52	52	52	-22	19	28
2	Элемент 2	100				100	100			100	100	26	-32	53
3	Элемент 3		132		132			132	132			-100	59	76
4	Элемент 4	74	74	74		74				74	74	42	-59	45
5	Элемент 5				74	74	74	74		74	74			39
6	Элемент 6	132						132	132	132		59	-100	76
7	Элемент 7			174	174						174	-59	42	82
8	Элемент 8	74	74			74	74	74		74		42	-59	45
9	Элемент 9	100		100			100	100		100		68	-132	72
	Ср.кв.откл.	50	50	62	66	43	45	58	58	44	60	57	65	

Из матрицы знаний (таблица 7) видно, что различные элементы несут различное количество информации о том, что предъявлена некоторая цифра, т.е. для каждой цифры одни элементы более характерны, а другие менее характерны.

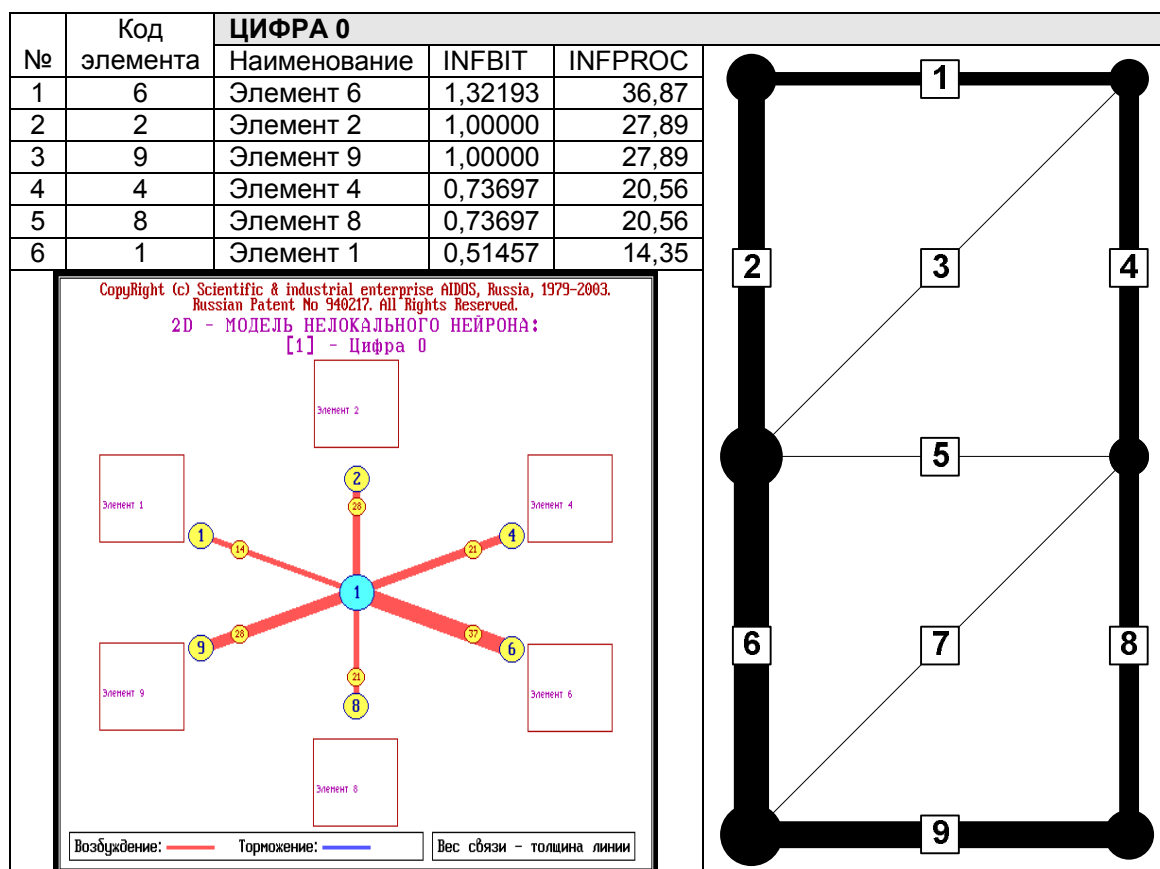
Например, для цифры «0» наиболее характерен 6-й элемент, за ним идут 2-й и 9-й элементы, а затем 4-й и 8-й, и самым нехарактерным для этой цифры является 1-й элемент.

Подобная информация представляется в системе «Эйдос» во многих различных формах, в частности в виде таблиц информационных портретов классов и рисунков нелокальных нейронов (таблица 8).

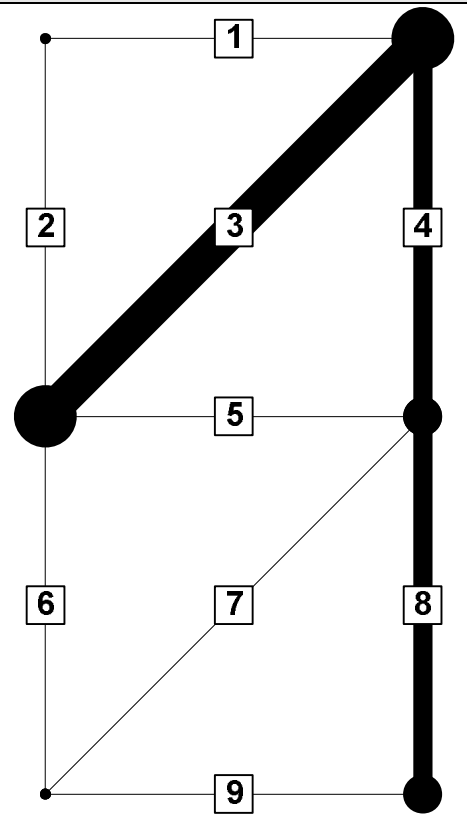
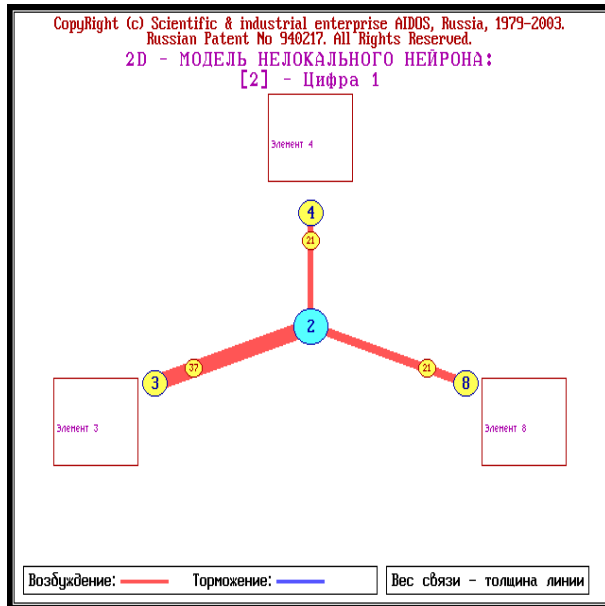
Предлагается визуализировать в изображениях значимость элементов путем их отображения цветом и толщиной линии, соответствующими значимости:

- *черный или красный цвет означает положительное количество информации;*
- *синий цвет означает отрицательное количество информации;*
- *толщина линии соответствует модулю количества информации.*

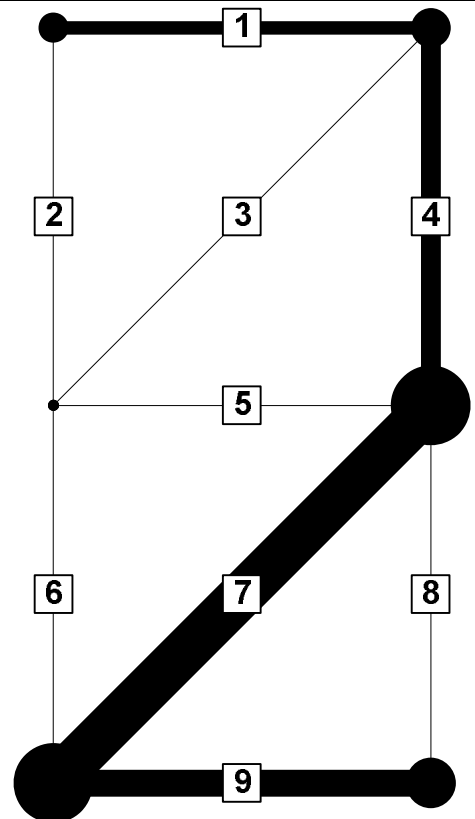
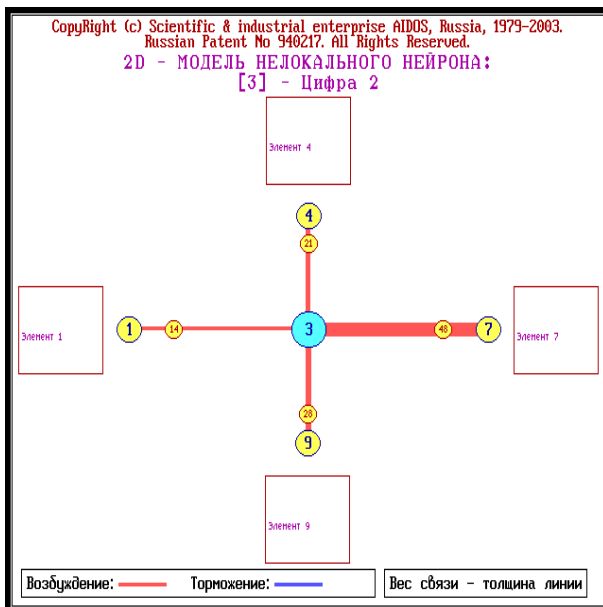
Таблица 8 – ИНФОРМАЦИОННЫЕ ПОРТРЕТЫ КЛАССОВ, НЕЛОКАЛЬНЫЕ НЕЙРОНЫ И ИЗОБРАЖЕНИЯ ЦИФР С ОТОБРАЖЕНИЕМ ЗНАЧИМОСТИ ЭЛЕМЕНТОВ ЦВЕТОМ И ТОЛЩИНОЙ ЛИНИИ

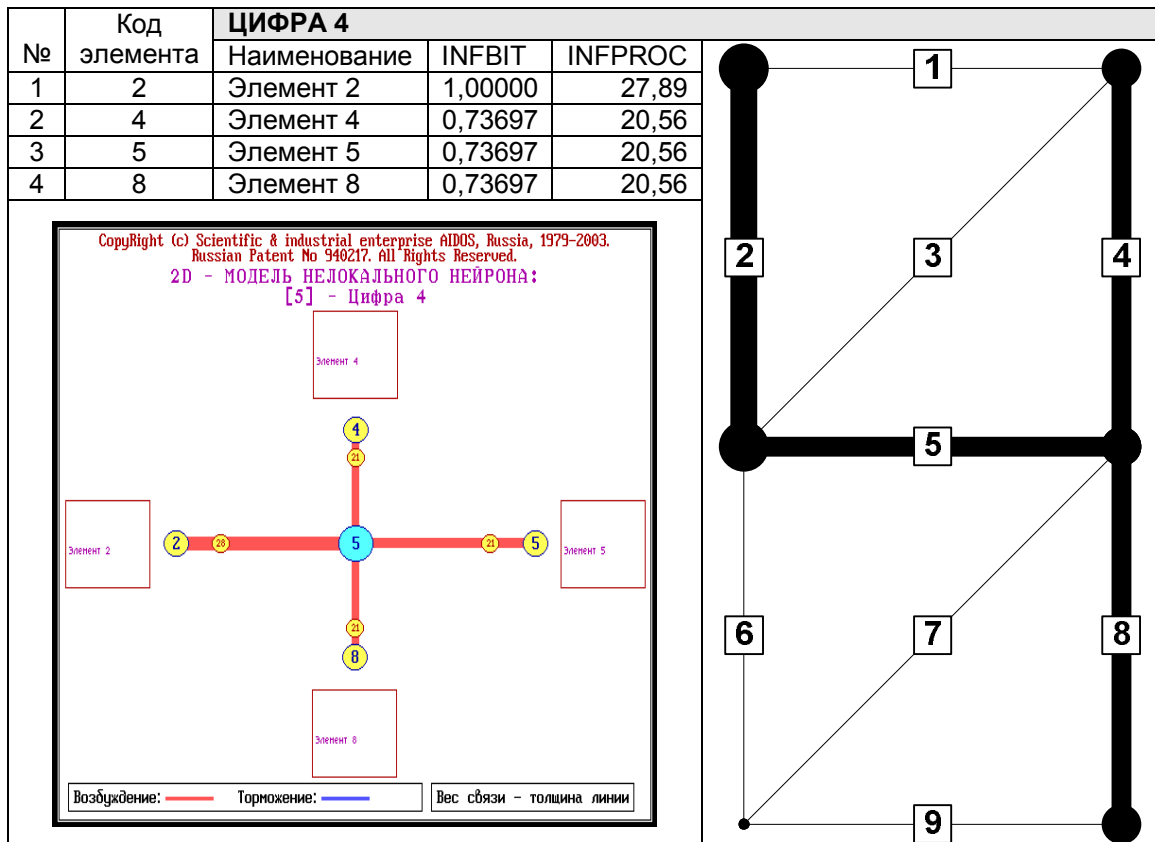
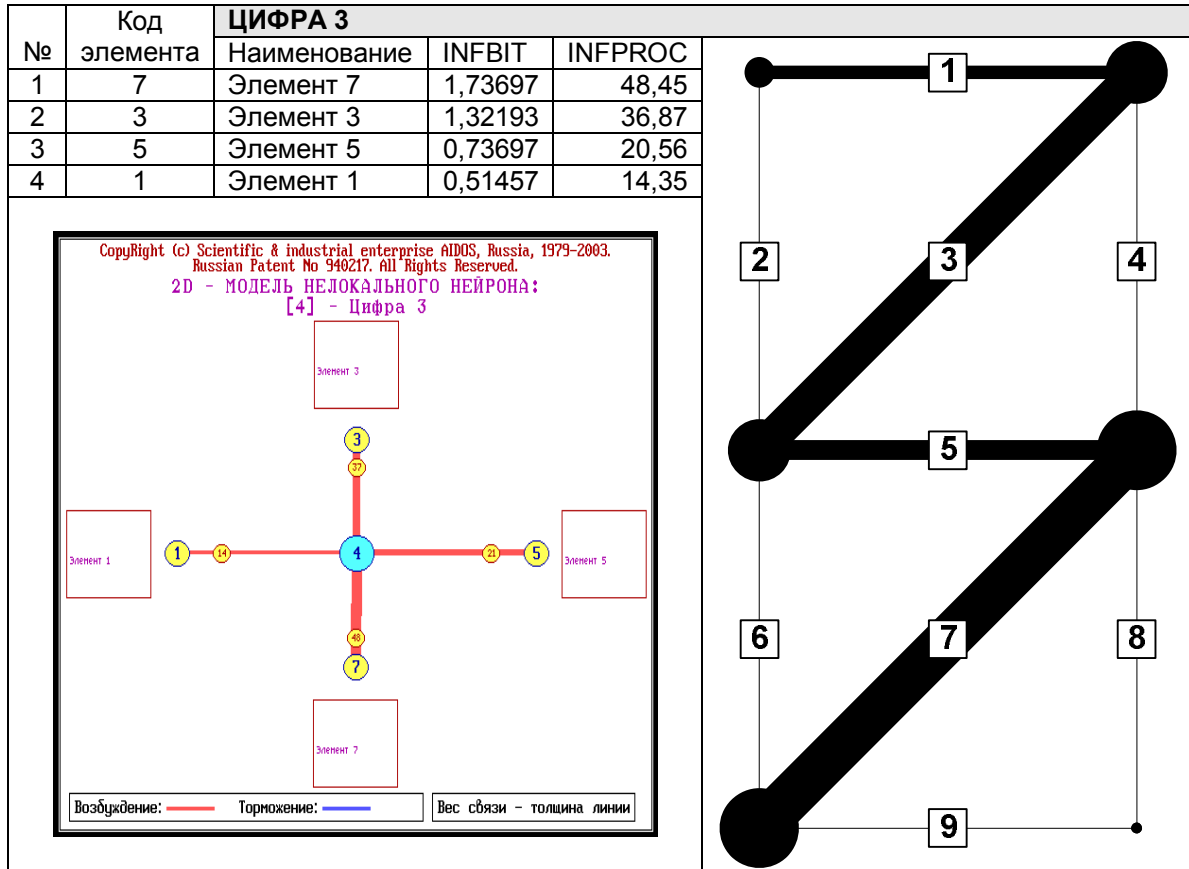


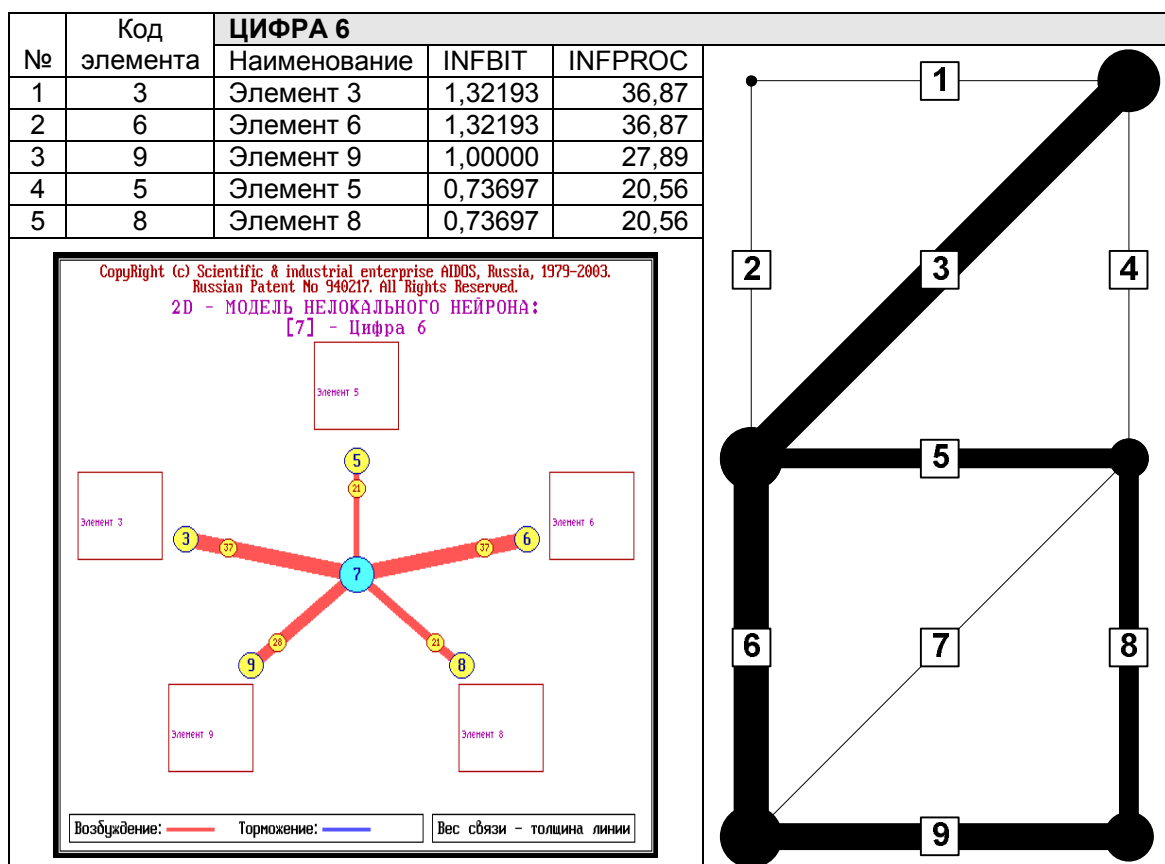
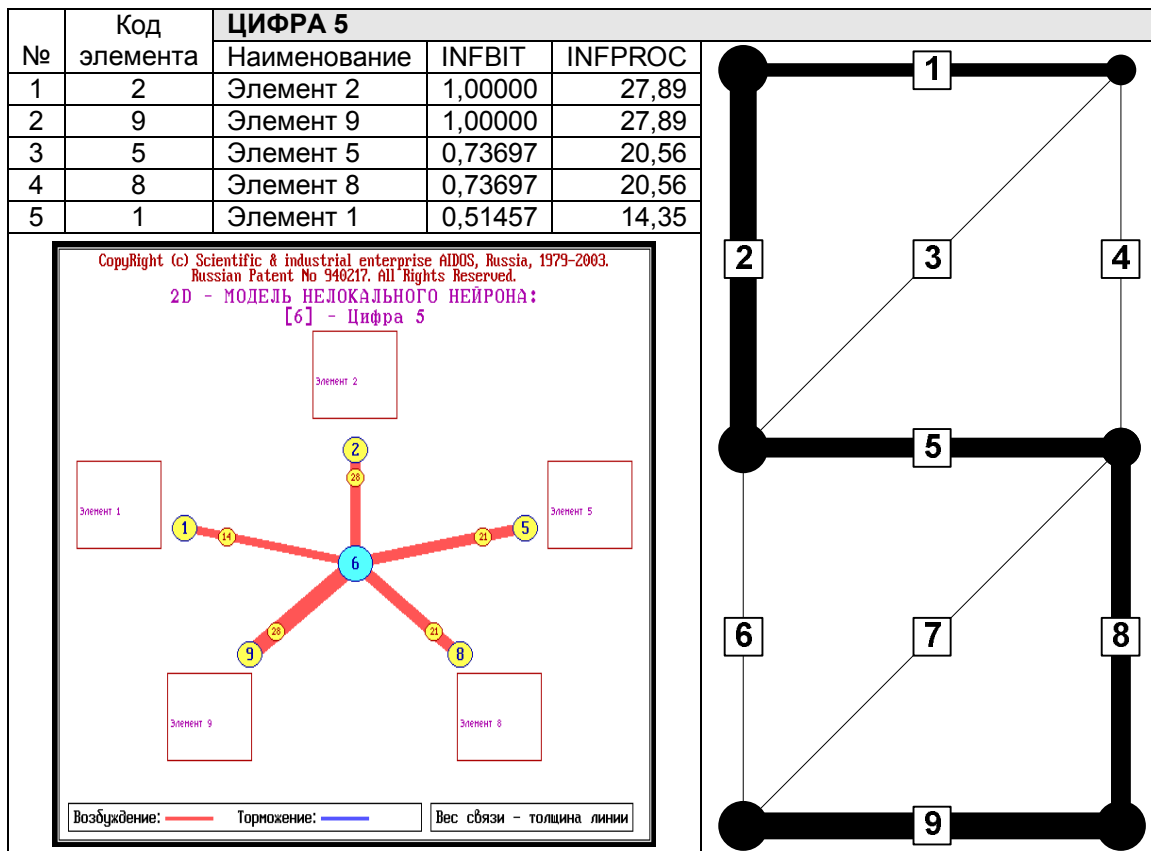
№	Код элемента	ЦИФРА 1		
		Наименование	INFBIT	INFPROC
1	3	Элемент 3	1,32193	36,87
2	4	Элемент 4	0,73697	20,56
3	8	Элемент 8	0,73697	20,56

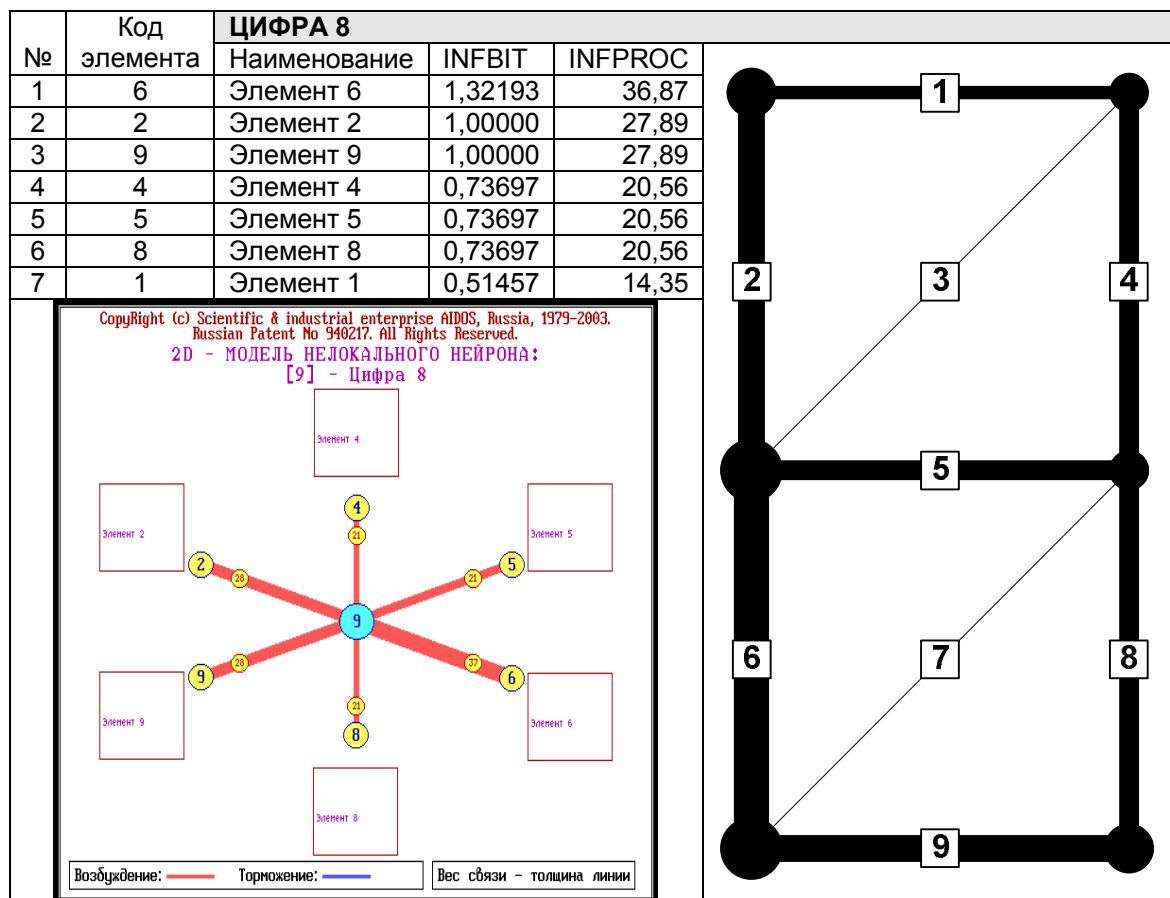
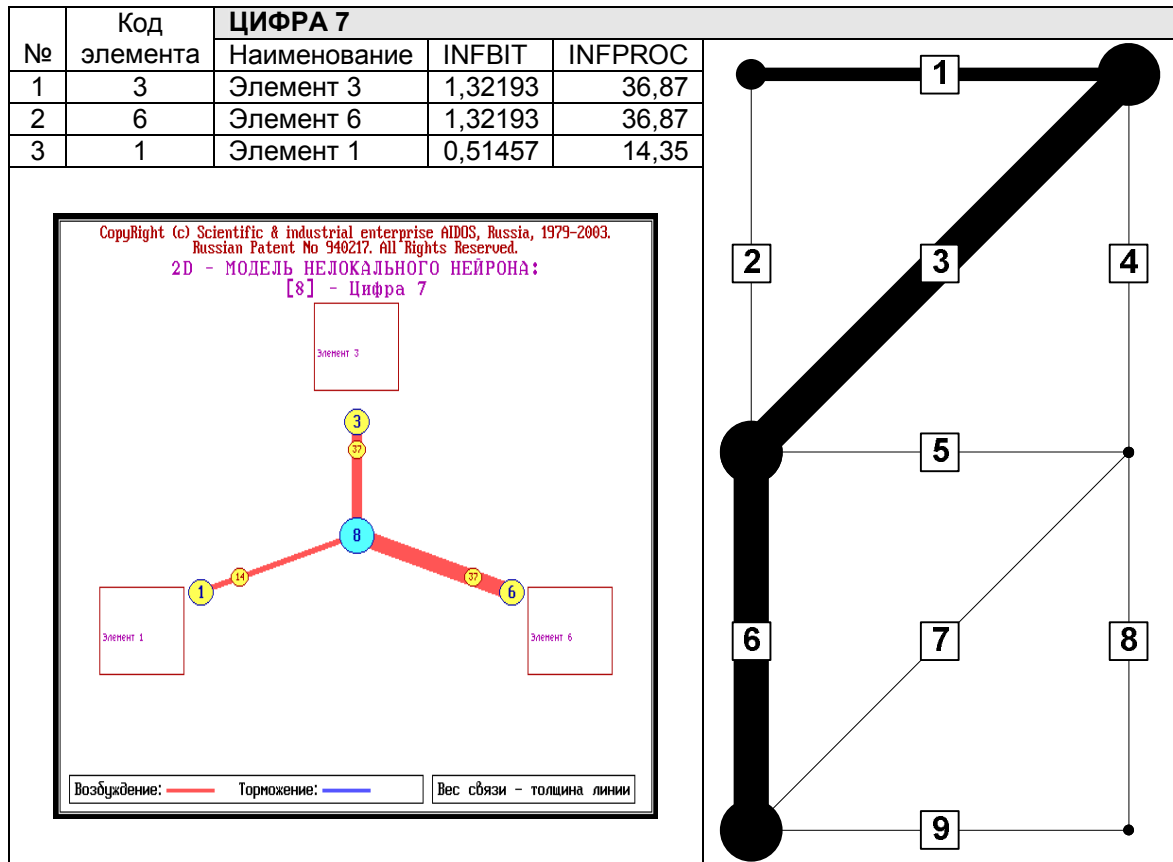


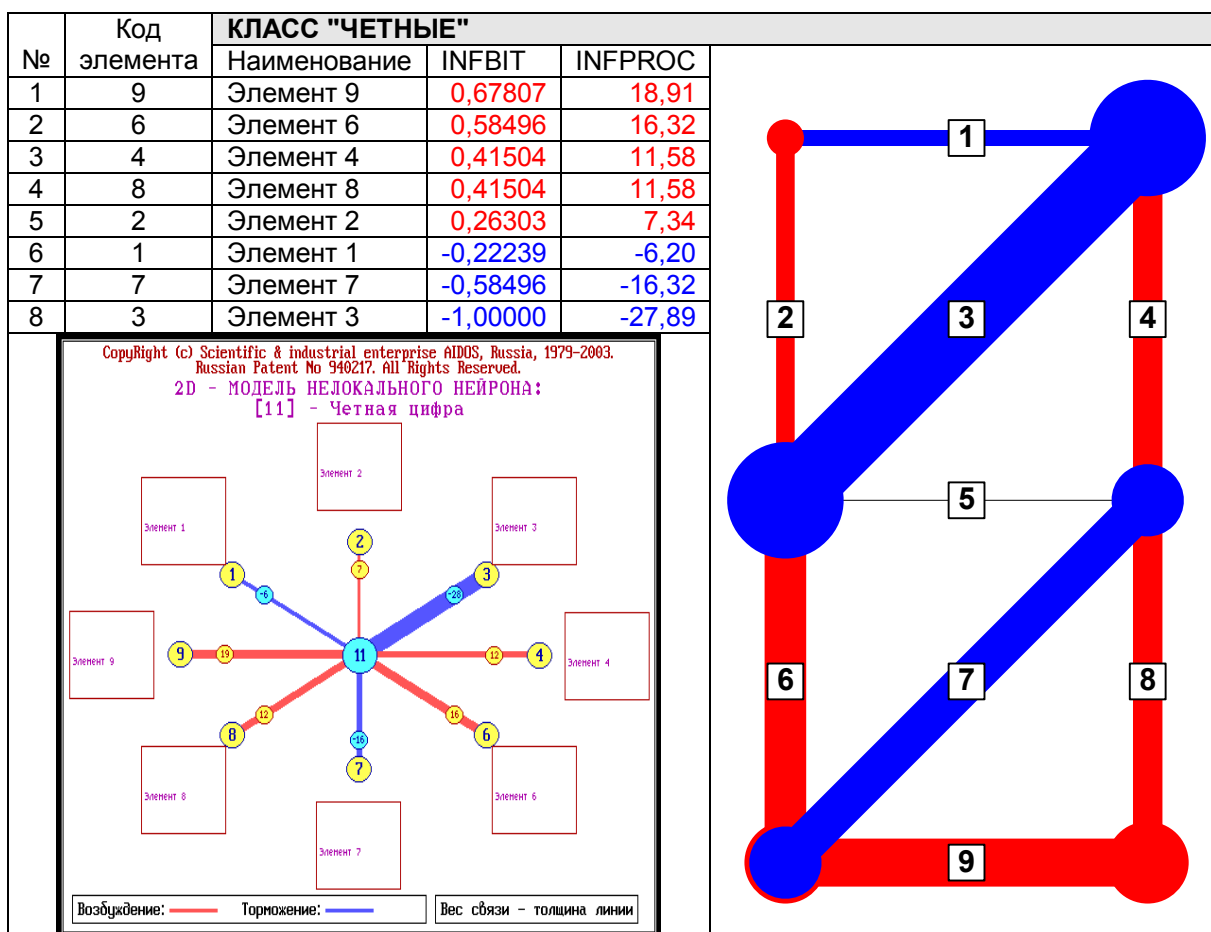
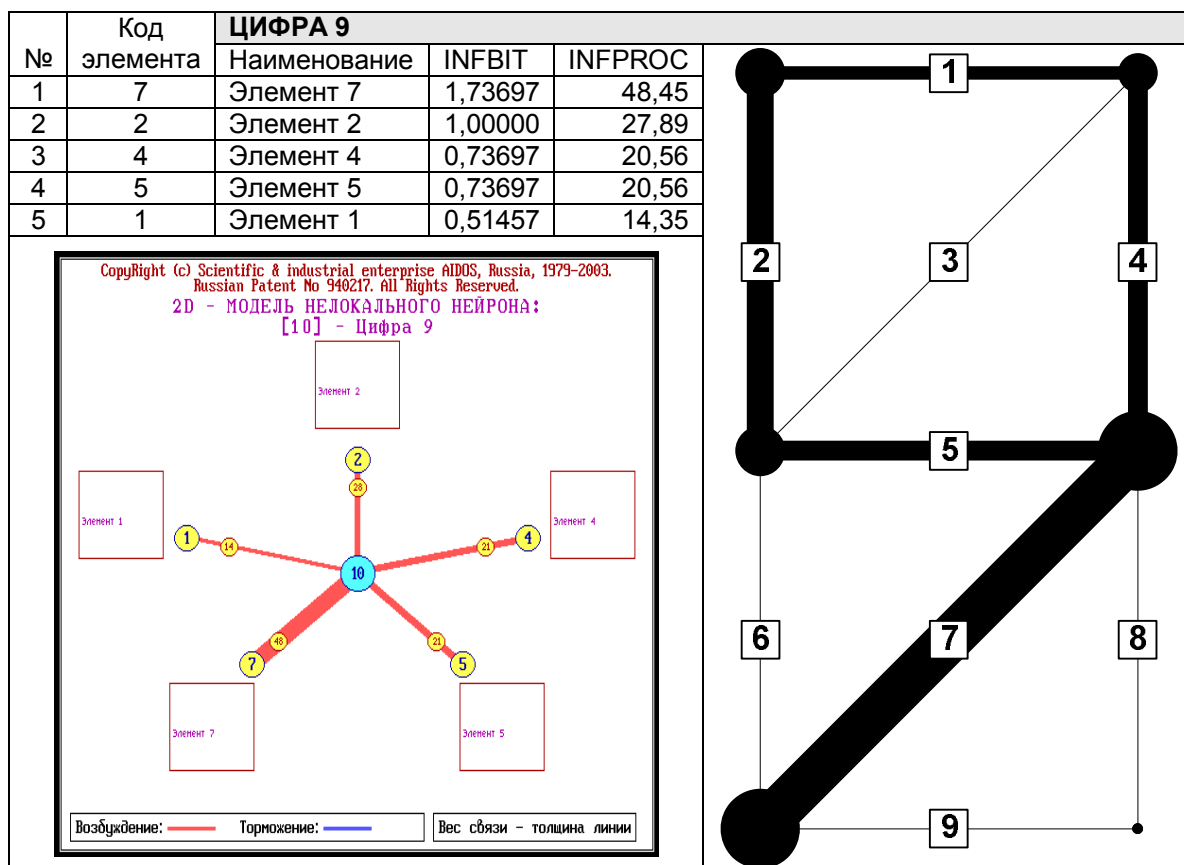
№	Код элемента	ЦИФРА 2		
		Наименование	INFBIT	INFPROC
1	7	Элемент 7	1,73697	48,45
2	9	Элемент 9	1,00000	27,89
3	4	Элемент 4	0,73697	20,56
4	1	Элемент 1	0,51457	14,35

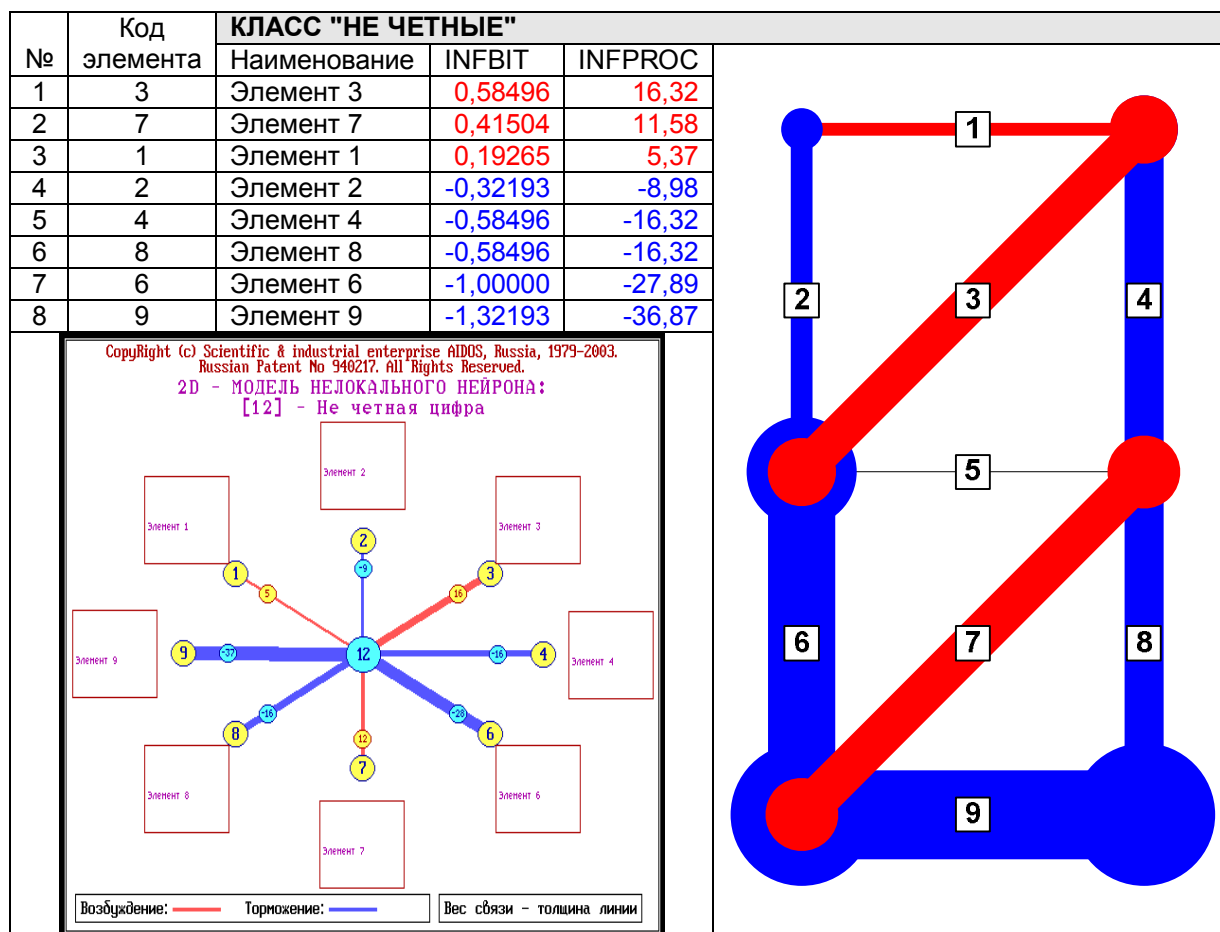












Приведем изображения десятичных цифр, используемые для написания почтовых индексов, с указанием значимости элементов, из которых они состоят, с помощью толщины линии или условно говоря ее «нажима» (рисунок 3).

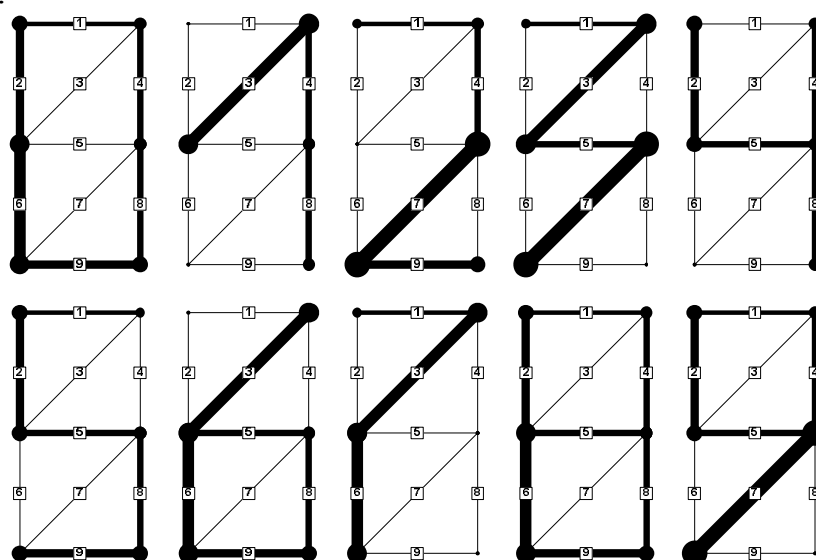


Рисунок 3. Изображения десятичных цифр, используемые для написания почтовых индексов, с указанием значимости элементов, из которых они состоят, с помощью толщины линии

Подобные изображения, в которых наиболее характерные их элементы каким-то образом выделены, причем пропорционально степени их характерности, принято называть *шаржами*. Обычно в искусстве выделение осуществляется с помощью размера, но можно это делать и с помощью яркости соответствующих элементов изображения, т.е. по сути *силы нажима* или *толщины линии*, как это и сделано в данной статье. Представляет интерес на наш взгляд и несколько неожиданный результат, состоящий в том, что оказывается в исследованном шрифте вертикальные и горизонтальные элементы цифр ассоциированы, содержат больше информации о принадлежности к четным цифрам, а наклонные – с нечетным.

Приведенная технология обеспечивает синтез *шаржей* любых типов изображений, например, таких как папиллярные узоры, радужная оболочка глаза, почерк [6] и фотороботы [7] лиц, сгруппированных по их полу, возрасту, профессии, социальному статусу, уровню образования, степени успешности тех или иных видов деятельности, например учебным или профессиональным достижениям по различным дисциплинам и циклам дисциплин, риску невозврата кредита, риску совершения ДТП и сумме страховых выплат КАСКО и ОСАГО и т.п. и т.д. Аналогичным образом можно обобщать изображения автомобилей, различных видов растений и пород животных, а также любые другие изображения, например предназначенные для восприятия элементы компьютерного интерфейса и дорожные знаки.

Обратимся к таблице 7, в которой проранжируем элементы в порядке убывания среднеквадратичного отклонения содержащегося в них количества информации о принадлежности изображений, включающих эти элементы, к классам цифр, а также к классам «четные» и «не четные». В результате получим (таблица 9):

Таблица 9 – СПИСОК ГРАФИЧЕСКИХ ЭЛЕМЕНТОВ, РАНЖИРОВАННЫЙ В ПОРЯДКЕ ЗНАЧИМОСТИ

№	Наименование	Значимость
1	Элемент 7	82
2	Элемент 3	76
3	Элемент 6	76
4	Элемент 9	72
5	Элемент 2	53
6	Элемент 4	45
7	Элемент 8	45
8	Элемент 5	39
9	Элемент 1	28

Из таблицы 9 видно, что 7-й элемент практически в 3 раза более значим, чем 1-й. Если бы элементов было больше, то малозначимые элементы без особого ущерба для адекватности модели вполне можно было бы удалить из нее. Точно также из образов цифр без особого ущерба можно уда-

лить малозначимые элементы. Эта операция и называется *абстрагированием*.

Сравнение друг с другом обобщенных образов классов (цифр), т.е. классификация, осуществляется системой «Эйдос». В результате формируется матрица сходства классов, т.е. изображений цифр (таблица 10):

Таблица 10 – МАТРИЦА СХОДСТВА КЛАССОВ

KOD	1	2	3	4	5	6	7	8	9	10	11	12
1	100,0	-28,7	-22,3	-86,8	11,2	25,8	11,8	6,5	86,9	-41,2	83,9	-85,4
2	-28,7	100,0	-28,4	15,6	5,2	-39,5	30,6	33,6	-45,0	-38,8	-39,2	28,4
3	-22,3	-28,4	100,0	41,9	-41,3	-22,9	-44,4	-39,9	-39,7	58,3	-13,9	7,2
4	-86,8	15,6	41,9	100,0	-43,6	-46,8	-5,2	16,4	-88,8	49,1	-90,4	83,4
5	11,2	5,2	-41,3	-43,6	100,0	43,3	-37,2	-56,8	30,8	19,3	34,7	-11,2
6	25,8	-39,5	-22,9	-46,8	43,3	100,0	-8,8	-55,4	42,4	-13,4	44,1	-35,5
7	11,8	30,6	-44,4	-5,2	-37,2	-8,8	100,0	63,8	19,2	-77,2	4,7	-27,2
8	6,5	33,6	-39,9	16,4	-56,8	-55,4	63,8	100,0	-5,4	-50,4	-32,8	15,8
9	86,9	-45,0	-39,7	-88,8	30,8	42,4	19,2	-5,4	100,0	-38,6	91,4	-85,9
10	-41,2	-38,8	58,3	49,1	19,3	-13,4	-77,2	-50,4	-38,6	100,0	-31,2	46,6
11	83,9	-39,2	-13,9	-90,4	34,7	44,1	4,7	-32,8	91,4	-31,2	100,0	-94,4
12	-85,4	28,4	7,2	83,4	-11,2	-35,5	-27,2	15,8	-85,9	46,6	-94,4	100,0

Фрагменты это матрицы сходства могут быть отображены средствами системы «Эйдос» в форме семантических сетей (рисунок 4):

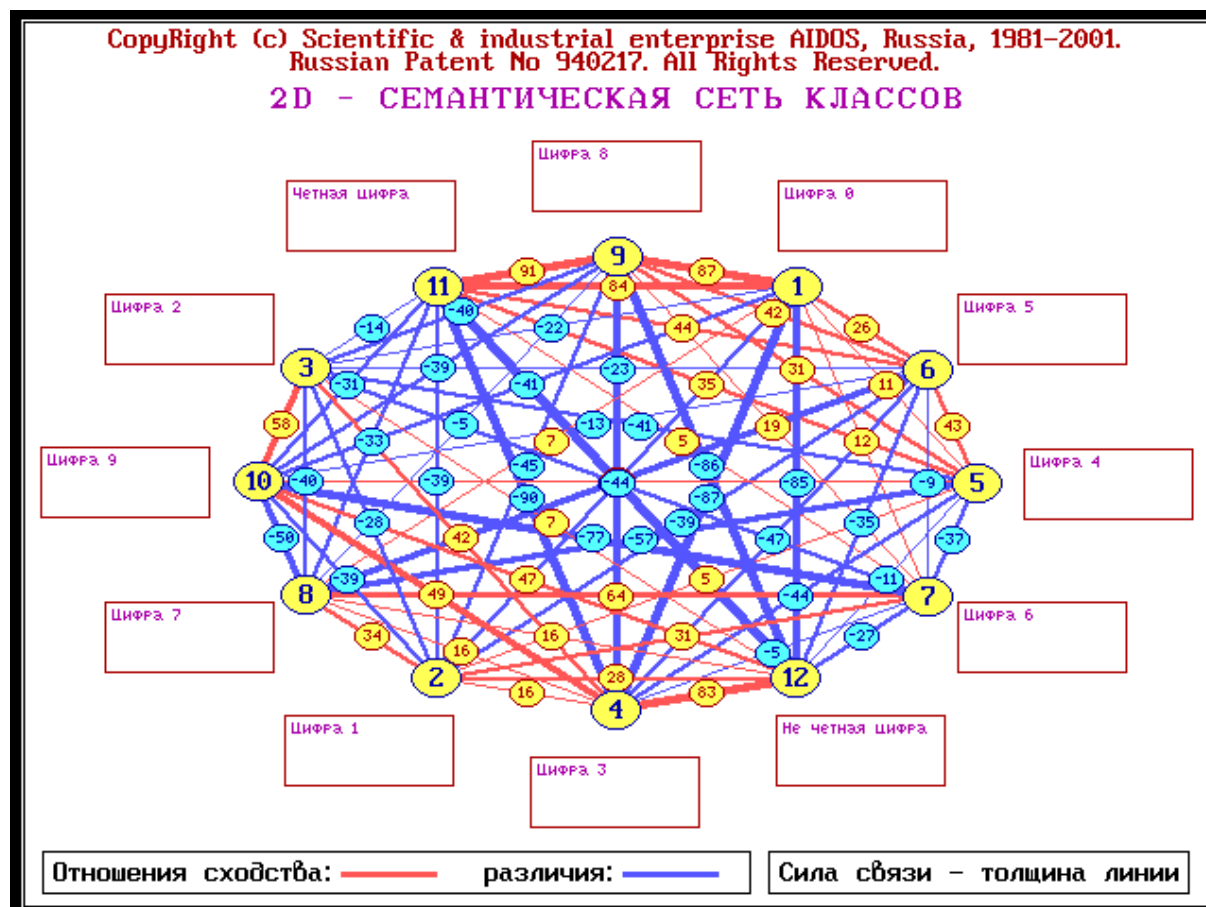


Рисунок 4. Семантическая сеть классов, отображающая в наглядной графической форме степень сходства и различия обобщенных образов классов (изображений)

Из рисунка 4 видно, что изображения цифр группируются в кластерах вокруг обобщенных образов четных и нечетных цифр, представляющих собой полюса конструкта (таблица 11):

Таблица 11 – КОНСТРУКТ «Не четная цифра – четная цифра»

№	Код класса	Наименование класса	Уровень Сходства (%)
1	12	Не четная цифра	100,00
2	4	Цифра 3	83,43
3	10	Цифра 9	46,59
4	2	Цифра 1	28,38
5	8	Цифра 7	15,84
6	3	Цифра 2	7,17
7	5	Цифра 4	-11,21
8	7	Цифра 6	-27,23
9	6	Цифра 5	-35,49
10	1	Цифра 0	-85,39
11	9	Цифра 8	-85,91
12	11	Четная цифра	-94,44

Обращает на себя внимание, что цифра «2» больше похожа на обобщенный образ нечетных цифр, а цифра «5» – на обобщенный образ четных цифр. На основе проведенного анализа можно даже сказать, что причина этого в том, что в изображении цифры «2» наиболее характерным является 7-й элемент, характерный именно для нечетных цифр, а в состав изображения цифры «5» входят элементы 2, 8 и 9, характерные для четных цифр.

Идентификация конкретных изображений цифр с их обобщенными образами осуществляется следующим образом:

– подсчитывается какое суммарное количество информации содержится в системе элементов данной конкретной цифры о ее принадлежности к каждому из обобщенных образов классов, сформированных в модели;

– классы ранжируются в порядке убывания суммарного количества информации о принадлежности к ним, содержащегося в систем признаков конкретной цифры;

– считается, что цифра относится к тому классу, о принадлежности к которому в ее системе признаков содержится максимальное количество информации.

Идентификация осуществляется в 4-й подсистеме системы «Эйдос». В результате формируются экранные формы карточек результатов идентификации и соответствующие выходные формы (рисунок 5):

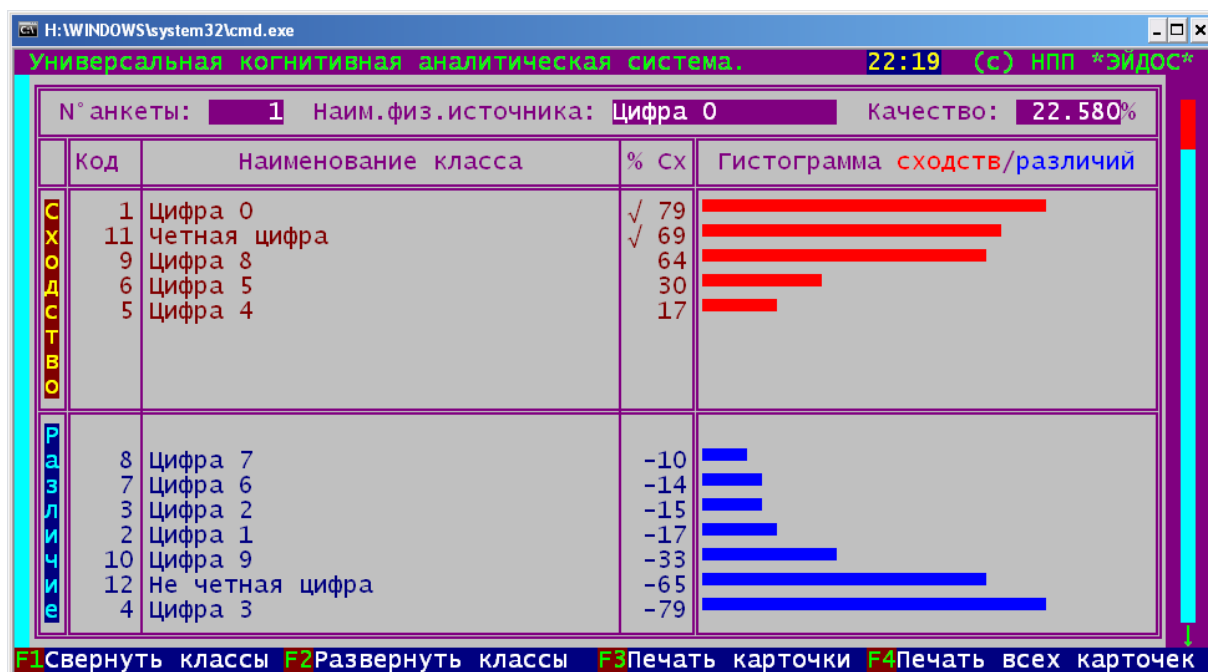


Рисунок 5. Экранная форма карточки идентификации цифры «0»

На рисунке 6 приведена экранная форма результатов идентификации изображений конкретных цифр с обобщенным образом «Не четная цифра».

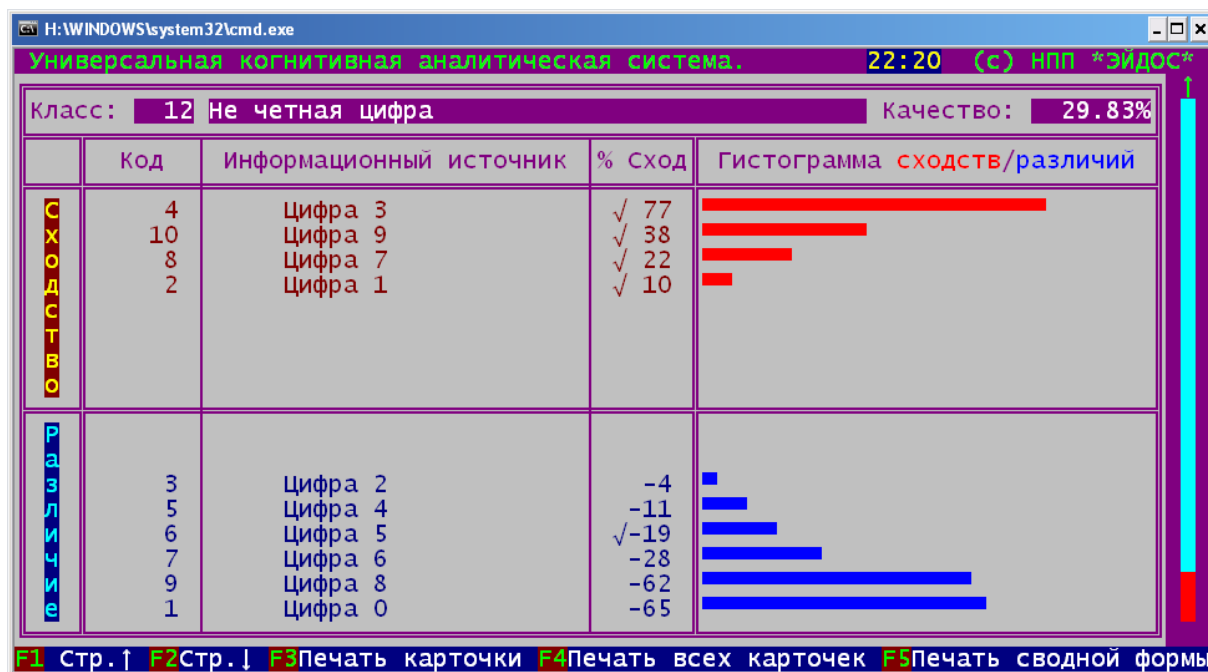


Рисунок 6. Экранная форма результатов идентификации изображений конкретных цифр с обобщенным образом «Не четная цифра»

Обобщение результатов идентификации позволяет измерить степень достоверности созданной модели (таблица 12):

Таблица 12 – ВЫХОДНАЯ ФОРМА ПО РЕЗУЛЬТАТАМ ИЗМЕРЕНИЯ ДОСТОВЕРНОСТИ МОДЕЛИ

ИЗМЕРЕНИЕ АДЕКВАТНОСТИ (ДИФФЕРЕНЦИАЛЬНОЙ И ИНТЕГРАЛЬНОЙ ВАЛИДНОСТИ) СЕМАНТИЧЕСКОЙ ИНФОРМАЦИОННОЙ МОДЕЛИ

Всего физических анкет: 10 (100% для п.15)
 Всего логических анкет: 20

4. Средняя достоверность идентификации логических анкет с учетом сходства : 23.473%
5. Среднее сходство логических анкет, правильно отнесенных к классу : 12.487%
6. Среднее сходство логических анкет, ошибочно не отнесенных к классу : 0.474%
7. Среднее сходство логических анкет, ошибочно отнесенных к классу : 4.874%
8. Среднее сходство логических анкет, правильно не отнесенных к классу : 16.333%
9. Средняя достоверность идентификации логических анкет с учетом кол-ва : 54.000%
10. Среднее количество физич-х анкет, действительно относящихся к классу: 3.000 (100% для п.11 и п.12)
 Среднее количество физич-х анкет, действительно не относящихся к классу: 7.000 (100% для п.13 и п.14)
 Всего физических анкет: 10.000 (100% для п.15)
11. Среднее количество и % лог-их анкет, правильно отнесенных к классу: 2.750, т.е. 91.667%
12. Среднее количество и % лог-их анкет, ошибочно не отнесенных к классу: 0.250, т.е. 8.333% (Ошибка 1-го рода)
13. Среднее количество и % лог-их анкет, ошибочно отнесенных к классу: 2.050, т.е. 29.286% (Ошибка 2-го рода)
14. Среднее количество и % лог-их анкет, правильно не отнесенных к классу: 4.950, т.е. 70.714%
15. Средневзвешенная вероятность случайного угадывания принадлежности объекта к классу (%): 30.000
16. Средневзвешенная эффективность применения модели по сравнению со случ. угадыванием (раз): 5.900
17. Обобщенная достоверность модели $(D1+D2)/2$: 81.190%. Обобщенная ошибка $(E1+E2)/2$: 18.810%

21-02-09 22:23:03

г.Краснодар

N п/п	Код класса	Наименование класса	Достов. идентиф. лог.анк. с уч.количества эвр.крит	Кол-во лог.анк. действ-но относящихся к классу	Количество логических анкет правильно или ошибочно отнесенных или не отнесенных к классу				Вероятн. случайного угадывания (%) =NLA/NFA	Эффективн. модели по срав. со случ. угадыв. (раз)
					Правильн. отнесен.	Ошибочно не отнес.	Ошибочно отнесен.	Правильн. не отнес.		
1	2	3	9	10	11	12	13	14	15	16
1	1	Цифра 0	0.0	1	1	0	5	4	10.000	10.000
2	2	Цифра 1	20.0	1	1	0	4	5	10.000	10.000
3	3	Цифра 2	60.0	1	1	0	2	7	10.000	10.000
4	4	Цифра 3	40.0	1	1	0	3	6	10.000	10.000
5	5	Цифра 4	0.0	1	1	0	5	4	10.000	10.000
6	6	Цифра 5	0.0	1	1	0	5	4	10.000	10.000
7	7	Цифра 6	60.0	1	1	0	2	7	10.000	10.000
8	8	Цифра 7	40.0	1	1	0	3	6	10.000	10.000
9	9	Цифра 8	20.0	1	1	0	4	5	10.000	10.000
10	10	Цифра 9	40.0	1	1	0	3	6	10.000	10.000
11	11	Четная цифра	80.0	5	5	0	1	4	50.000	2.000
12	12	Не четная цифра	80.0	5	4	1	0	5	50.000	1.600
Ср.взв.значения			54.0	3.0	2.8	0.3	2.1	5.0	30.000	5.900

Универсальная когнитивная аналитическая система

НПП «ЭЙДОС»

ФОРМУЛЫ РАСЧЕТА ПОКАЗАТЕЛЕЙ ДИФФЕРЕНЦИАЛЬНОЙ ВАЛИДНОСТИ (ПО КЛАССАМ):

$$C04[k] = C05[k] - C06[k] - C07[k] + C08[k]$$

$$C09[k] = (C11[k] - C12[k] - C13[k] + C14[k]) / (C11[k] + C12[k] + C13[k] + C14[k]) * 100$$

$$C10[k] = C11[k] + C12[k]$$

$$C15[k] = C10[k] / Nfiz * 100$$

$$C16[k] = C09[k] / C15[k]$$

где k – класс (соответствует строке)

где Nfiz – суммарное количество физических анкет (объектов) в распознаваемой выборке

ФОРМУЛЫ РАСЧЕТА ПОКАЗАТЕЛЕЙ ИНТЕГРАЛЬНОЙ ВАЛИДНОСТИ (СРЕДНЕВЗВЕШЕННОЕ ПО ВСЕМ КЛАССАМ):

$$Ci = \text{СУММА_по_k}(Ci[k] * C10[k]) / NLog$$

где i = { 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 }

где NLog = СУММА_по_k(C10[k]) – суммарное количество логических анкет в распознаваемой выборке

ПРИМЕЧАНИЕ: учтены только результаты идентификации с модулем сходства не менее: 0+

Таким образом, в статье рассмотрено применение системно-когнитивного анализа, его математической модели – системной теории информации (СТИ) и программного инструментария – системы «Эйдос» для синтеза обобщенных изображений классов, их абстрагирования, классификации обобщенных изображений (кластеры и конструкты) сравнения конкретных изображений с обобщенными образами (идентификация).

Рассмотрение проведено на конкретном примере изображений цифр шрифта, используемого для написания почтовых индексов, но примененные при этом методы являются универсальными и могут быть использованы и при обработке других самых разнообразных изображений.

Необходимо также отметить, что в рассмотренном примере все цифры верно отнесены к классам, к которым они относятся, причем каждая цифра является наиболее похожей из всех именно на класс, к которому она действительно принадлежит, т.е. ошибка неидентификации равна нулю, однако к ним отнесены и другие цифры, которые к ним не относятся, т.е. *есть ошибка ложной идентификации*.

Учитывая это можно предположить, что актуальной является задача *разработки таких шрифтов и других знаков, предназначенных для визуального восприятия (например элементов компьютерного интерфейса, дорожных знаков и т.п.), начертания которых минимизировали бы ошибки как 1-го, так и 2-го рода, т.е. были бы наиболее легко и безошибочно воспринимаемы как человеком, так и системами машинного зрения. Это имеет значение как для минимизации затрат различных видов вычислительных и других ресурсов и времени, так и для улучшения восприятия затухающих текстов, чтения в условиях зашумленности или погрешностей самого аппарата восприятия.*

Обоснованной выглядит гипотеза, что шрифт, в котором наиболее информативные элементы специально выделены, например яркостью или толщиной элементов как в представленном на рисунке 3, является более легким для восприятия, чем исходный шрифт с одинаковой толщиной графических элементов, представленный на рисунке 1. Однако экспериментальная проверка этой гипотезы требует проведения инженерно-психологических исследований и не входит в задачу данной статьи.

По-видимому, предложенный в статье подход может быть применен и для разработки эффективных *рекламных* стимульных материалов, а также в *Ψ-технологиях*². Материалы данной статьи могут быть использованы также в качестве основы для лабораторных и курсовых работ по дисциплине «Интеллектуальные информационные системы», а также дипломных работ по специальностям «Прикладная информатика (по отраслям)».

² См., например, И.В.Смирнов: <http://www.psycor.ru/main.php?smir>

Литература

1. Луценко Е.В. Теоретические основы и технология адаптивного семантического анализа в поддержке принятия решений (на примере универсальной автоматизированной системы распознавания образов "ЭЙДОС-5.1"). Монография (научное издание). – Краснодар: КЮИ МВД РФ, 1996. – 280с.
2. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): Монография (научное издание). – Краснодар: КубГАУ. 2002. – 605 с.
3. Луценко Е.В. Интеллектуальные информационные системы: Учебное пособие для студентов специальности 351400 "Прикладная информатика (по отраслям)". – Краснодар: КубГАУ. 2004. – 633 с.
4. Луценко Е.В. Интеллектуальные информационные системы: Учебное пособие для студентов специальности "Прикладная информатика (по областям)" и другим экономическим специальностям. 2-е изд., перераб. и доп.– Краснодар: КубГАУ, 2006. – 615 с.
5. Луценко Е.В. Математическая сущность системной теории информации (СТИ) (Системное обобщение формулы Больцмана-Найквиста-Хартли, синтез семантической теории информации Харкевича и теории информации Шеннона) / Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №08(42). – Шифр Информрегистра: 0420800012\0114. – Режим доступа: <http://ej.kubagro.ru/2008/08/pdf/04.pdf>
6. Луценко Е.В. Прогнозирование учебных достижений студентов на основе особенностей их почерка с применением системно-когнитивного анализа / Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2006. – №04(20). – Шифр Информрегистра: 0420600012\0083. – Режим доступа: <http://ej.kubagro.ru/2006/04/pdf/27.pdf>
7. Луценко Е.В. Возможности прогнозирования учебных достижений студентов на основе АСК-анализа их имеджевых фотороботов / Е.В. Луценко // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2004. – №02(4). – Режим доступа: <http://ej.kubagro.ru/2004/02/pdf/13.pdf>

Примечание:

Для обеспечения доступа читателей к этим и другим работам они размещены в Internet по адресу: <http://lc.kubagro.ru/aidos/>