



Е. В. Луценко, Л. П. Трошин

***КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ
В АМПЕЛОГРАФИИ***

Монография

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГБОУ ВО «Кубанский государственный аграрный университет
имени И. Т. Трубилина»

Е. В. Луценко, Л. П. Трошин

**КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ
В АМПЕЛОГРАФИИ**

МОНОГРАФИЯ

Краснодар
КубГАУ
2024

УДК 634.84(075.8)

ББК 42.36

Л 86

Рецензенты:

А. К. Раджабов – зав. кафедрой виноградарства Московской сельскохозяйственной академии имени К. А. Тимирязева,
д-р с.-х. наук, профессор, Москва;
П.П. Радчевский – зав. кафедрой виноградарства Кубанского
госагроуниверситета, к. с.-х. н., доцент ВАК РФ,
лауреат Госпремии РФ, Ялта.

Луценко Е. В., Трошин Л. П.

Л86 Когнитивное моделирование в ампелографии : монография /
Е. В. Луценко, Л. П. Трошин. – Краснодар : КубГАУ, 2024. – 153 с.

ISBN 978-5-93856-883-9

В монографии рассматривается феномен полиморфизма в ампелографии, характеризующий разнообразие внутри сортов винограда, и его значение для формирования адаптивных ампелоценозов. Описаны применения автоматизированного системно-когнитивного анализа (АСК-анализа) и интеллектуальной системы «Эйдос» для задач оцифровки и анализа листьев винограда. Предложен новый подход к оцифровке с использованием полярной системы координат и стандартизации изображений. Рассматриваются методы выделения прототипа формы листа сорта или клона, независимого от случайных воздействий среды. Представленный подход позволяет идентифицировать сорта и клоны, определять степень их сходства и различия, что обеспечивает значительный вклад в развитие виноградарства и ампелографии.

УДК 634.84(075.8)

ББК 42.36

© Луценко Е. В.,

Трошин Л. П., 2024

© ФГБОУ ВО «Кубанский
государственный аграрный
университет имени
И. Т. Трубилина», 2024

DOI: [10.13140/RG.2.2.28319.06566](https://doi.org/10.13140/RG.2.2.28319.06566)

ISBN 978-5-93856-883-9

*Монография подготовлена в рамках
реализации программы развития
Кубанского ГАУ «Приоритет 2030»
(стратегический проект «Генетика
и селекция в растениеводстве»).*

ОГЛАВЛЕНИЕ

| | |
|--|-----|
| ПРЕДИСЛОВИЕ | 5 |
| 1. ВВЕДЕНИЕ В АМПЕЛОМЕТРИЮ..... | 7 |
| 2. РЕШЕНИЕ ЗАДАЧ АМПЕЛОГРАФИИ С ПРИМЕНЕНИЕМ АСК-АНАЛИЗА ИЗОБРАЖЕНИЙ ЛИСТЬЕВ ПО ИХ ВНЕШНИМ КОНТУРАМ (ОБОБЩЕНИЕ, АБСТРАГИРОВАНИЕ, КЛАССИФИКАЦИЯ И ИДЕНТИФИКАЦИЯ) | 11 |
| 3. КОЛИЧЕСТВЕННОЕ ИЗМЕРЕНИЕ СХОДСТВА-РАЗЛИЧИЯ КЛОНОВ ВИНОГРАДА ПО КОНТУРАМ ЛИСТЬЕВ С ПРИМЕНЕНИЕМ АСК-АНАЛИЗА И СИСТЕМЫ «ЭЙДОС»..... | 60 |
| 4. ПРИМЕНЕНИЕ ТЕОРИИ ИНФОРМАЦИИ И КОГНИТИВНЫХ ТЕХНОЛОГИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ГЕНЕТИКИ (НА ПРИМЕРЕ ВЫЧИСЛЕНИЯ КОЛИЧЕСТВА ИНФОРМАЦИИ В ГЕНАХ О ПРИЗНАКАХ И СВОЙСТВАХ РАЗЛИЧНЫХ АВТОХТОННЫХ СОРТОВ ВИНОГРАДА) | 83 |
| 5. БИОМЕТРИЧЕСКАЯ ОЦЕНКА ПОЛИМОРФИЗМА СОРТОГРУПП ВИНОГРАДА ПИНО И РИСЛИНГ ПО МОРФОЛОГИЧЕСКИМ ПРИЗНАКАМ ЛИСТЬЕВ СРЕДНЕГО ЯРУСА КРОНЫ* | 138 |
| ОБЩИЕ ВЫВОДЫ | 151 |

ПРЕДИСЛОВИЕ

Протекавшие в последние годы перестройка отечественной науки и гармонизация образования приводят к современной интеграции близкородственных биологических и сельскохозяйственных направлений, как-то виноградарства, виноделия, биоэкологии винограда, ампелографии, клоновой и генетической селекций, общей и молекулярной генетики, увологии, ампелометрии и ампелологии, что на сегодня выразилось в яркой вспышке большого количества созданных и интродуцированных генотипов винограда – генетов и ренетов...

В этом глубинном процессе интеграции усилий ученых, педагогов и производителей активное участие принимают селекционеры и питомниководы России из разных НИИ и опытных станций не только нашего Отечества, но и зарубежных. Особо следует отметить коллективный труд исследователей под координирующим руководством некогда хорошо работавшей секции виноградарства Российской сельхозакадемии, объединявшей усилия работников Всероссийского НИИВиВ имени Я.И. Потапенко, Национального НИИВиВ «Магарач», Северо-Кавказского ЗНИИСиВ, Анапской ЗОСВиВ, Крымской ОСС ВИРа, Дагестанской ОС и Кубанского госагроуниверситета, на основе творческого сотрудничества – зарубежного Афинского института виноградарства (Греция). В интеграции науки и производства заметную роль в России ныне играют не только производственные коллективы ООО «Южная», ООО «Собер Баш», «Фанагория», «Победа», «Кубанский виноград» и «Мысхако», но и многочисленные фермеры, а также любители виноградной лозы.

В данной монографии отражены результаты интеграционных процессов науки, образования и производства, венцом которых явились созданные многочисленные сорта и клоны винограда. И чтобы история ампелографии и

ампелометрии не допускала «ошибок», каждый генотип снабжен своим авторским коллективом, авторской карточкой.

Основой монографии являются следующие работы:

<http://ej.kubagro.ru/2009/08/pdf/01.pdf>,

<http://ej.kubagro.ru/2015/08/pdf/64.pdf>,

<http://ej.kubagro.ru/2016/02/pdf/77.pdf>,

<http://ej.kubagro.ru/2016/07/pdf/03.pdf>.

Авторы монографии выражают искренние признательность и благодарность ректорату университета, сотрудникам кафедры виноградарства и кафедры компьютерных технологий за помощь в проведении университетских научно-исследовательских работ и в компоновке представленного труда.

...ниже: какие гиганты мысли, рук и дел –
мыслители, ваятели и мастера!



Рафаэль Санти. Афинская школа. 1510-1511

1. ВВЕДЕНИЕ В АМПЕЛОМЕТРИЮ

Как сказано в первом томе мировой известности трехтомной «Энциклопедии виноградарства» (Кишинев, 1967), «АМПЕЛО... (от греч. *ampelos* – виноград), первая составная часть сложных слов, обозначающая «относящийся к винограду»... АМПЕЛОМЕТРИЯ (от *ампело...* и *...метрия*), раздел ампелографии, включающий специзмерения количественных показателей изучаемых частей виноградного растения или его сортовых признаков и определение степени их варьирования с применением методов биометрии. Возникновению А. способствовали работы французского ампелографа Раваза (1902), который разработал законченную систему измерений пластинки виноградного листа, углов нервации, глубины вырезок, высоты зубчиков и т.д.

Именно этому древнейшему культурному растению принадлежит лидирующая роль быть первым среди магических, обладающих уникальной способностью делать людей радостными и счастливыми...Сами убедитесь, покопайтесь в древних рукописях и проверьте...

Среди огромнейше-многообразного звукосочетания гласно-несогласных букв врывается в человеческое сознание гипнотический всему миру известный русскоязычный комплекс-конгломерат... именно о нем много лет назад было сказано следующее: «Вначале слышно слово – слово твердое, звучное, удивительно интригующее и сложное. В нем отзвук развивающейся земной цивилизации. В едином звукосочетании совершенно разных двух слов – вино (*vin* – капелька Вселенной, в переводе с санскрита), и град (в переводе со старославянского «город») образуется уникальное – виноград. Это непревзойденное по красоте формы и ценности содержания чудо природы вызывает у каждого человека гамму чувств и ассоциативных впечатлений, возвышающих настроение,

обогащающих нравственное состояние и потому осветляющих и облегчающих жизнь» [1].

У непознавших в реальности эту материальную вещь – виноград – сам термин своим дуалистическим словосочетанием направляет интерес вначале к своим свойствам и признакам, а затем, по мере углубленного общения, к познанию таинственных историй, и оказывается, виноградные лозы известны человечеству с доисторических времен, более того – из-за них происходили многие столкновения не только внутри родов, племен и народов, но и между ними. Познавшие же прелесть зрелых ягод или продуктов их переработки навсегда оставались их «пленниками», ибо виноград – растение магическое, завораживающее и увлекающее в свои до сих пор необъяснимые тайны навсегда.

По эзотерическим данным, в виноградном растении заключена особая магическая сила (заметим, крупные социально-политические потрясения всегда следуют за очередными кампаниями варварской борьбы с виноградниками – источниками одухотворяющих гроздей, соков и напитков). И лоза стонет: не трогайте меня! Не уничтожайте! В нас жизнь! И она, лоза, действительно одухотворяет нашу жизнь. А жизнь – это латинское *vita*. Именно поэтому во всех странах мира ученые ныне называют семейство виноградных лоз *Vitaceae*, выделяя его уникальную необыкновенность.

На всех этапах развития человеческого общества виноград вызывал к себе уважение изумрудной живой зеленью, высокой питательностью и великой животворной силой. С глубокой любовью его воспевали крупнейшие мастера слова – поэты и прозаики всех времен и цивилизованных народов.

И как можно не знать знаменитейшие слова Иисуса Христа (От Иоанна, глава 15):

1 Я есмь истинная виноградная лоза, а Отец Мой - виноградарь.

2 Всякую у Меня ветвь, не приносящую плода, Он отсекает; и всякую, приносящую плод, очищает, чтобы более принесла плода.

5 Я есмь лоза...

А сколько картин знаменитых и малоизвестных художников отражают его внешние особенности! В честь него написан балет, создано множество поговорок и пословиц.

Усиливающееся стремление к общению с внешне чудной и содержательно неповторимой плодоносящей виноградной лозой и к ее окультуриванию является признаком возрастающей цивилизованности общества. Те народы успешнее прогрессировали интеллектуально, которые ритуально употребляли строго ограниченные дозы вина. Трудно даже представить себе отсутствие названного растительного создания на Земле, без которого бы «наша кормилица-матушка» потеряла свою своеобразную привлекательность, ибо плоды-ягоды лианообразных виноградных лоз в свежем или переработанном виде – это более чем всегда желаемые лакомства, они – источник милых вдохновений, эйфории и чудодейственных последствий.

Всегда собранное множество разноформенных и неравновеликих, но чаще всего эффектных гроздей винограда, состоящих из искрящихся броской зрелостью, неповторимых по окраске ягод, навеивает мысль как бы о стремлении самой природы одухотворить и накормить каждого уравненным количеством лакомых единичных виноградинок – «бусинок жизни», так легко отделяемых от гармонично целостного «города ягод». И как примечательна эта всегдашняя возможность разделения грозди на равные части по числу желающих это сделать (каждому по веточке)! Но чтобы это воспринимать и ощущать, надо предварительно иметь такие эффектные грозди.

Клинописные и папирусные источники информации свидетельствуют, что «сын юга – виноград» издревле возделывался в Северной Африке («Винограда в долине Нила

также много, как воды в Ниле. У разных сортов свои свойства, свой цвет, и сортов этих много») и восточном Средиземноморье («легенда о чудесном винограде, одну гроздь которого с трудом могли нести два человека. Этот виноград рос на земле Ханаанской...») [2].

И из этих вторичных очагов происхождения культурного винограда (современное знание указывает на первичность Закавказья и Ближнего Востока как основных очагов возникновения виноградарства и виноделия) в течение многих тысячелетий велось распространение семян, одревесневших лоз и вегетирующих растений во все страны света, где происходили скрещивания с местными формами, обмены генетической информацией, мутационные процессы, и, прежде всего, по большим дорогам мировых цивилизаций. Такими путями и в ходе сознательно и бессознательно производимых отборов обогащался сортимент и образовалась масса сортов и форм. Развивающаяся наука ампелография на сегодня описала по развернутой программе лишь немногим более 20 тысяч сортов и форм из более 50 тысяч наименований [3].

И неудивительно, что культура винограда с оригинальными узорами роскошных побегов лианообразных кустов, наполненных нарядными гроздьями ароматных ягод, весьма привлекательна для наблюдений и опытов (попутно заметим – десять, не семь, «цветов радуги» окрасок ягод и более 600 биохимических компонентов – это ли не уникальная природная скоррелированность формы и содержания!) [4].

И перефразируя разноплановые слова виднейших ученых, напоминаем исследователям о математических подходах при изучении законов и закономерностей природы, ее сути и эволюции, в частности, виноградной лозы как природного объекта и самого сложного фрагмента – ампелографии с быстро развивающейся биометрически формализованной ампелометрией.

2. РЕШЕНИЕ ЗАДАЧ АМПЕЛОГРАФИИ С ПРИМЕНЕНИЕМ АСК-АНАЛИЗА ИЗОБРАЖЕНИЙ ЛИСТЬЕВ ПО ИХ ВНЕШНИМ КОНТУРАМ (ОБОБЩЕНИЕ, АБСТРАГИРОВАНИЕ, КЛАССИФИКАЦИЯ И ИДЕНТИФИКАЦИЯ)

В главе рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и программного инструментария – интеллектуальной системы «Эйдос» для решения ряда задач ампелографии: 1) оцифровка сканированных изображений листьев и создание их математических моделей; 2) формирование математических моделей конкретных листьев с применением теории информации; 3) формирование моделей обобщенных образов листьев различных сортов; 4) сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу; 5) количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов. Предлагается новый подход к оцифровке изображений листьев, основанный на использовании полярной системы координат, центра тяжести изображения и его внешнего контура. Перед оцифровкой изображений могут применяться их преобразования, стандартизирующие положение изображений, их размеры и угол поворота. Поэтому результаты оцифровки и АСК-анализа изображений могут быть инвариантны (независимы) относительно их положения, размеров и поворота. Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о сорте, включающее как информацию об истинной форме листа данного сорта (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным воздействием окружающей среды.

Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение сигнала об истинной форме листа каждого сорта на основе ряда зашумленных конкретных примеров листьев данного сорта. Таким образом, создается один образ формы листа каждого сорта, независимый от их конкретных реализаций, т.е. «Эйдос» этих изображений (в смысле Платона) - прототип или архетип (в смысле Юнга) изображений [21].

*Раз он в море закинул невод, -
Пришел невод с одною тиною.
Он в другой раз закинул невод,
Пришел невод с травой морскою.
В третий раз закинул он невод, -
Пришел невод с одною рыбкой,
С непростой рыбкой, - золотою.
/А.С.Пушкин/*

Данная глава может рассматриваться как продолжение серии работ [1-5], посвященных применению автоматизированного системно-когнитивного анализа (АСК-анализ) [12] и его программного инструментария – системы «Эйдос» [10, 11]¹ для интеллектуальной обработки изображений, т.е. для их оцифровки, создания моделей конкретных изображений, формирования обобщенных изображений на основе ряда конкретных, относящихся к одной категории (классу), абстрагирования, идентификации, классификации обобщенных изображений и решения ряда других задач.

В частности, в работе [3] описаны основы АСК-анализа изображений по их внешним контурам. В этих работах описаны возможности применения АСК-анализа для решения задачи *синтеза обобщенных изображений* на основе ряда конкретных примеров. 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 *абстрагирование обобщенных изображений*, 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 сильно отличающихся друг от друга кластеров, т.е. конструкторов. Можно также количественно оценивать степень сходства конкретных изображений с обобщенными, т.е. идентифицировать эти конкретные изображения.

Все это позволяет ставить и решать на практике ряд важных задач ампелографии:

¹ См. сайт автора АСК-анализа и системы «Эйдос»: <http://lc.kubagro.ru/>

- 1) оцифровка сканированных изображений листьев и создание их математических моделей;
- 2) формирование математических моделей конкретных листьев с применением теории информации;
- 3) формирование моделей обобщенных образов листьев различных сортов;
- 4) сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу;
- 5) количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов.

Предлагается новый подход к оцифровке изображений листьев винограда, основанный на использовании полярной системы координат, центра тяжести изображения и его внешнего контура. Перед оцифровкой изображений могут применяться их преобразования, стандартизирующие положение изображений, их размеры и угол поворота. Поэтому результаты оцифровки и АСК-анализа изображений могут быть инвариантны (независимы) относительно их положения, размеров и поворота. Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о сорте, включающее как информацию об истинной форме листа данного сорта (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным воздействием окружающей среды. Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение сигнала об истинной форме листа каждого сорта на основе ряда зашумленных конкретных примеров листьев данного сорта. Таким образом, создается один образ формы листа каждого сорта, независящий от их конкретных реализаций, т.е. «Эйдос»

этих изображений (в смысле Платона [6]) - прототип или архетип (в смысле Юнга [18]) изображений.

Рассмотрим на конкретном численном примере решение сформулированных выше задач ампелографии в АСК-анализе и системе «Эйдос-X++» версии от 24.10.2015 и выше. Отметим, что система «Эйдос» имеет программные интерфейсы с внешними источниками данных различных типов: текстовые данные в форме текстовых файлов и значений полей Excel-таблиц, табличными данными dbf, xls и xlsx форматов, графическими объектами, анализируемыми по пикселям и внешним контурам.

Для решения поставленных в главе задач используем программный интерфейс с изображениями, которые анализируются по их внешним контурам в полярной системе координат. Программная реализация данного интерфейса разработана Д.К. Бандык по алгоритму и постановке проф. Е.В. Луценко [3, 4].

Для этого скачаем систему «Эйдос» с сайта автора по ссылке: http://lc.kubagro.ru/aidos/_Aidos-X.htm и установим ее в соответствии с инструкцией на сайте.

Запишем в папку ..\AID_DATA\Inp_data\ сканированные изображения листьев (рисунок 1):

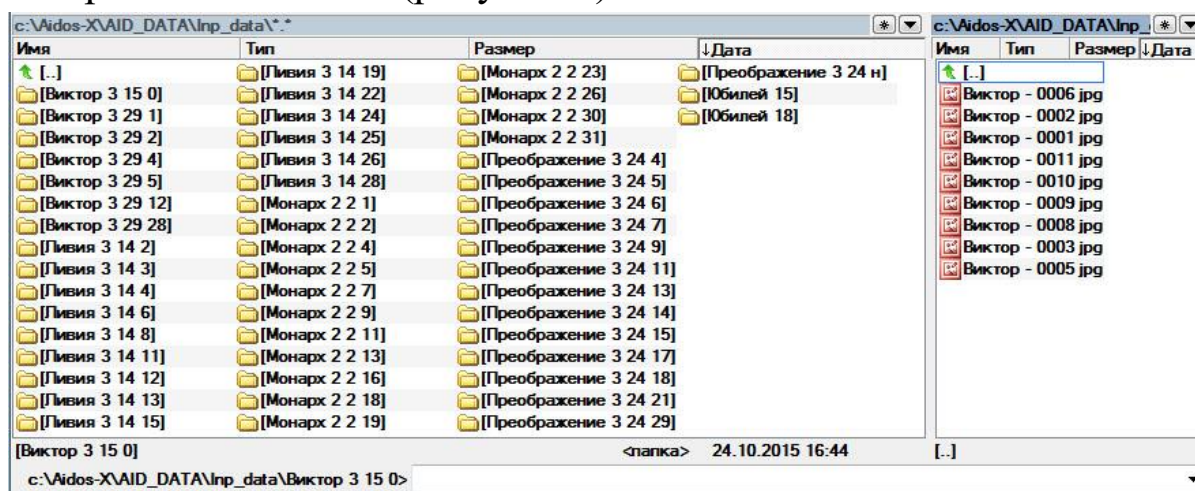


Рисунок 1. Экранная форма, поясняющая расположение и структуру исходных данных: сканированных изображений листьев различных сортов винограда

Изображения листьев помещены в папки, имена которых содержат информацию о сорте и месторасположении куста (слева). Внутри каждой папки (справа) содержатся файлы сканированных изображений листьев с одного куста. **Имена файлов состоят из имени сорта (класса) и номера листа на кусте и внутри папки** (номер реализации). В рассматриваемом примере в обучающей выборке используются изображения 451 листа.

Затем запустим режим: 2.3.2.4. Оцифровка изображений по их внешним контурам (рисунок 2):

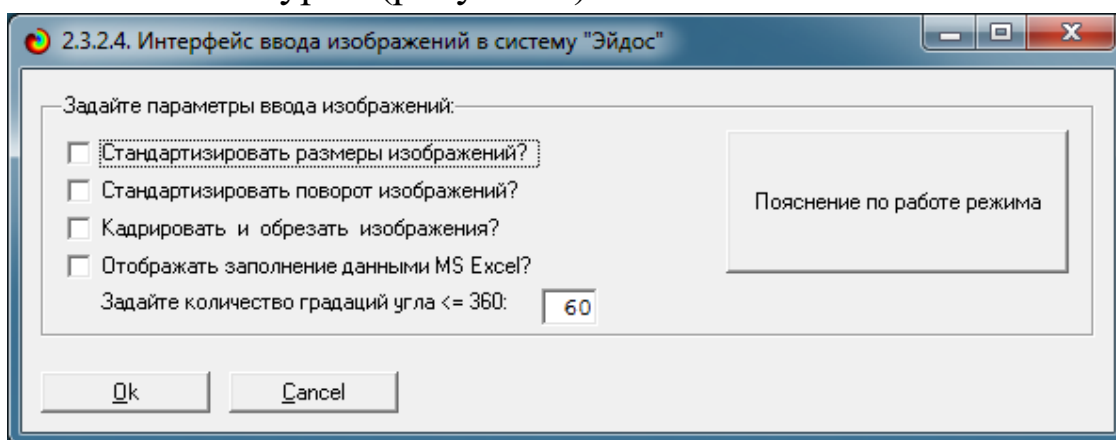


Рисунок 2. Начальная экранная форма режима 2.3.2.4. Оцифровка изображений по их внешним контурам

Вместо описания данного режима приведем на рисунке 3 Help данного режима:

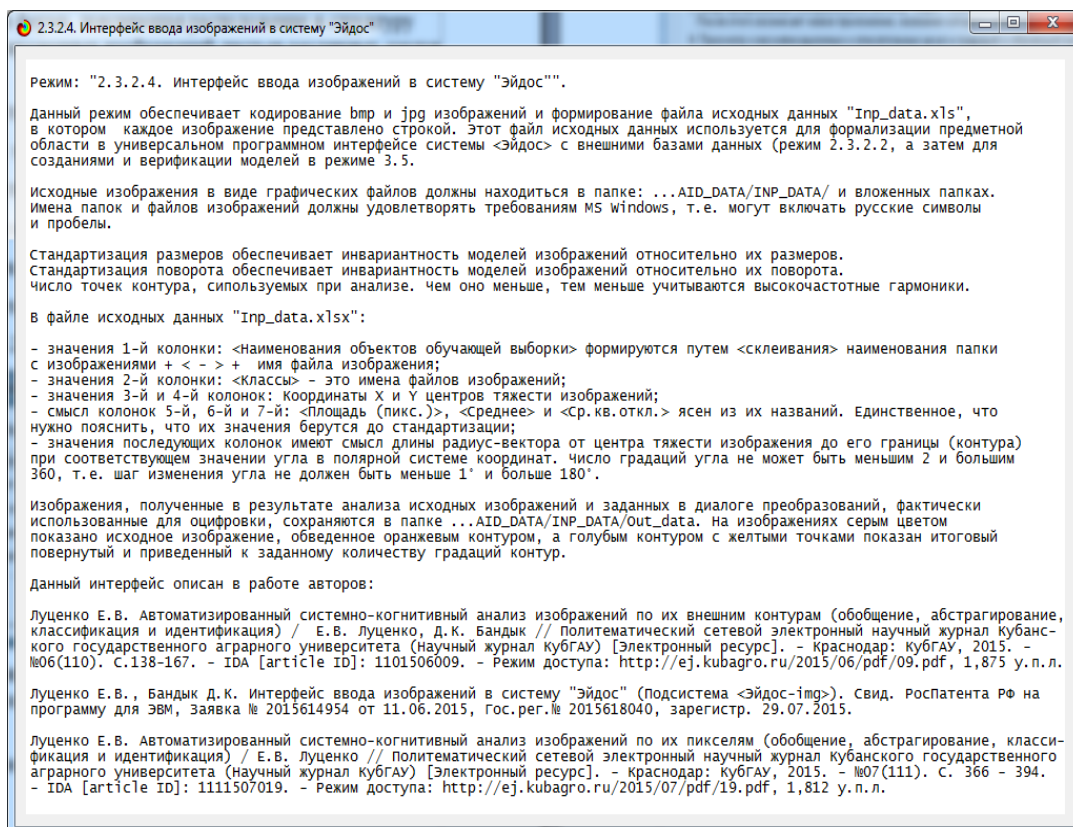


Рисунок 3. Help режима 2.3.2.4. Оцифровка изображений по их внешним контурам

В результате появляется главная экранная форма, отображающая процесс оцифровки сканов листьев, расположенных в виде файлов в указанных выше папках (рисунок 4):

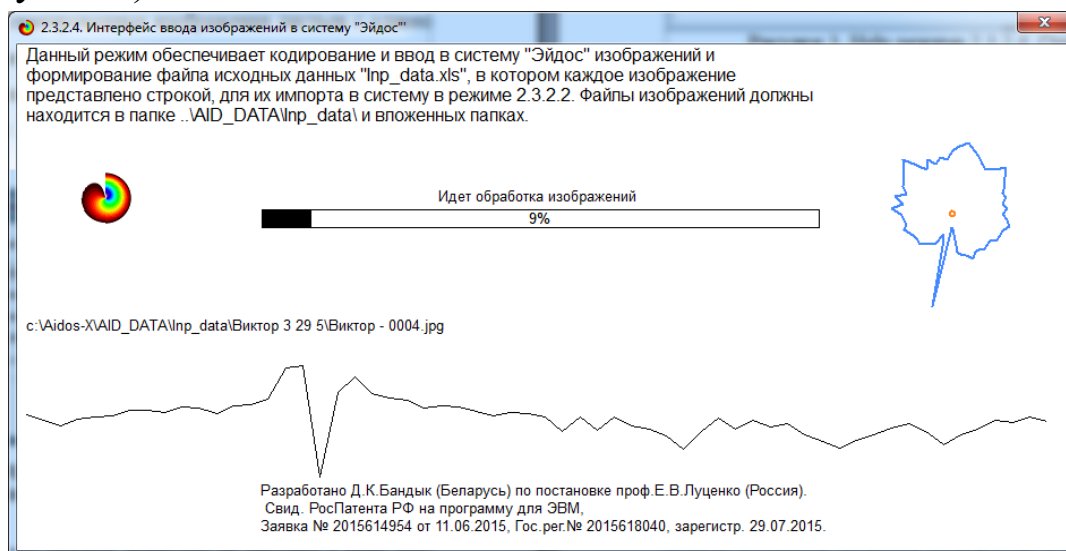


Рисунок 4. Экранная форма с отображением стадии процесса исполнения оцифровки изображений листьев по их внешним контурам

Данный режим: 1. Находит все поддиректории в папке: ..\AID_DATA\Inp_data\ и все графические файлы jpg и bmp в поддиректориях. 2. Находит контуры в этих графических файлах и их центры тяжести. 3. Записывает в папку: ..\AID_DATA\Out_data\ графические файлы, состоящие только из контуров с изображенными на них точками, которые были оцифрованы. **Необходимо особо отметить, что при этом используется полярная система координат с центром в центре тяжести изображения, а результатами оцифровки являются расстояния от центров тяжести изображений до точек их контура при различных углах поворота радиуса-вектора** (эта идея, а также математическая модель и алгоритм ее реализации предложены проф. Е.В. Луценко в 2014 и реализованы в модуле 2.3.2.4 и режиме 4.7 системы «Эйдос» в 2015 году. При этом структура поддиректорий и имена файлов в папках: \AID_DATA\Inp_data\ и ..\AID_DATA\Out_data\ совпадают. Пример контурного изображения листа приведен ниже на рисунке 5:

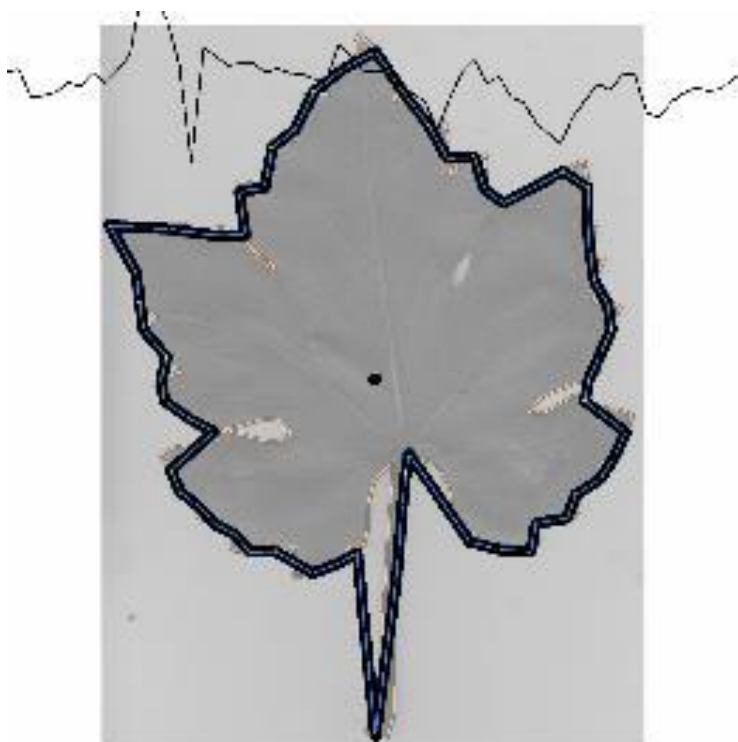


Рисунок 5. Пример контурного изображения листа:
с:\Aidos-X\AID_DATA\Out_data\Виктор 3 15 0\Виктор - 0010.jpg

4. Затем режим 2.3.2.4 формирует Excel-таблицу с именем: ..\AID_DATA\Inp_data\inp_data.xlsx, в которой содержатся результаты оцифровки изображений. Кроме того, он формирует таблицу: ..\AID_DATA\Inp_data\inp_data_avr.xlsx с усредненными данными по классам. Пояснение по структуре этой таблицы дано в Help режима 2.3.2.4 (рисунок 3). Структура этой таблицы полностью соответствует требованиям универсального программного интерфейса системы «Эйдос» с внешними базами данных (режим 2.3.2.2), которые приведены в Help этого режима и представлены на рисунке 6:

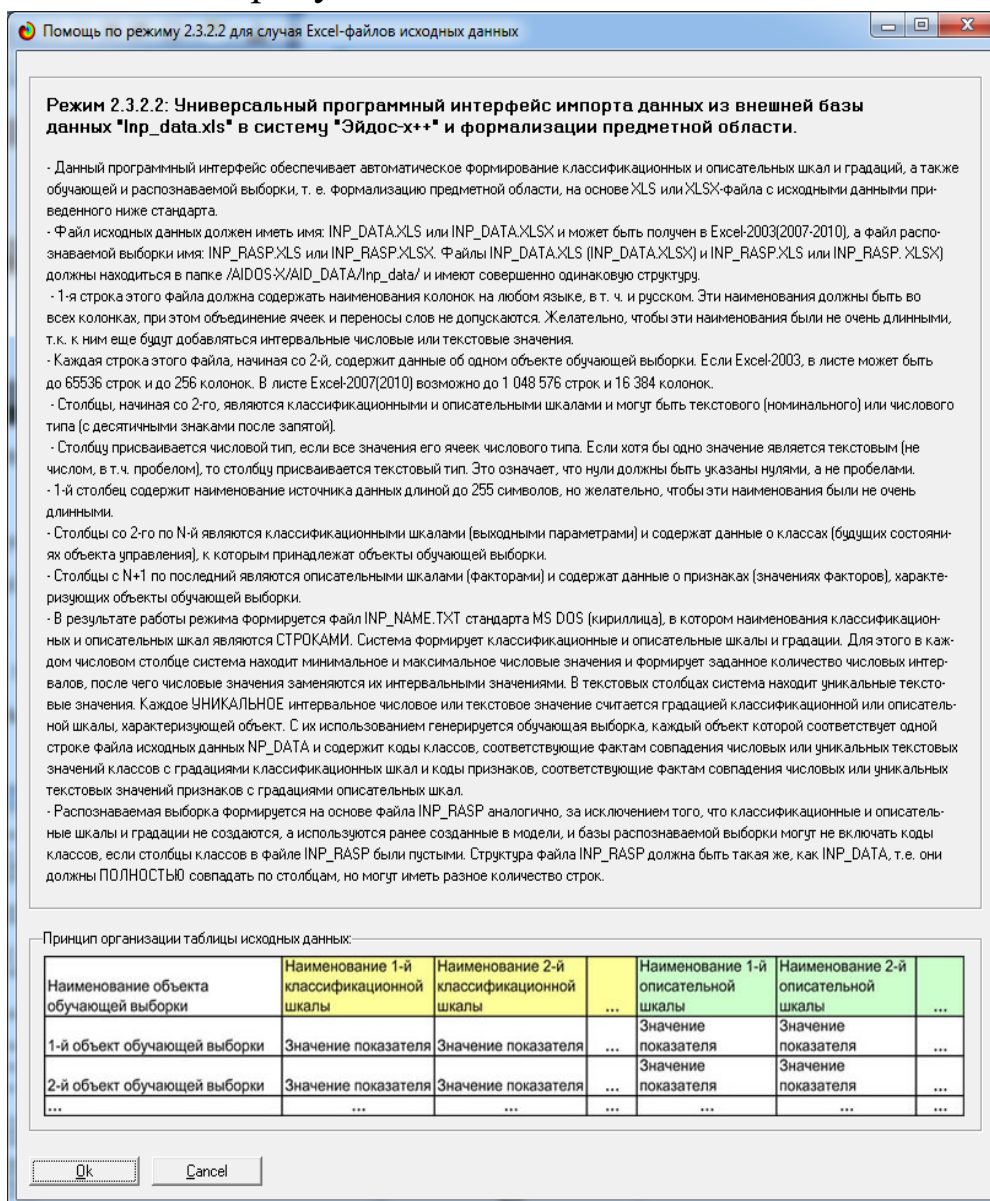


Рисунок 6. Help универсального программного интерфейса системы «Эйдос» с внешними базами данных (режим 2.3.2.2)

Сам этот программный интерфейс вызывается из главного меню системы «Эйдос» (режим 2.3.2.2), либо из режима: 4.7. АСК-анализ изображений. Главная экранная форма режима 2.3.2.2 приведена на рисунке 7, причем на ней показаны нужные параметры, которые формируются режимом 2.3.2.4 как параметры по умолчанию:

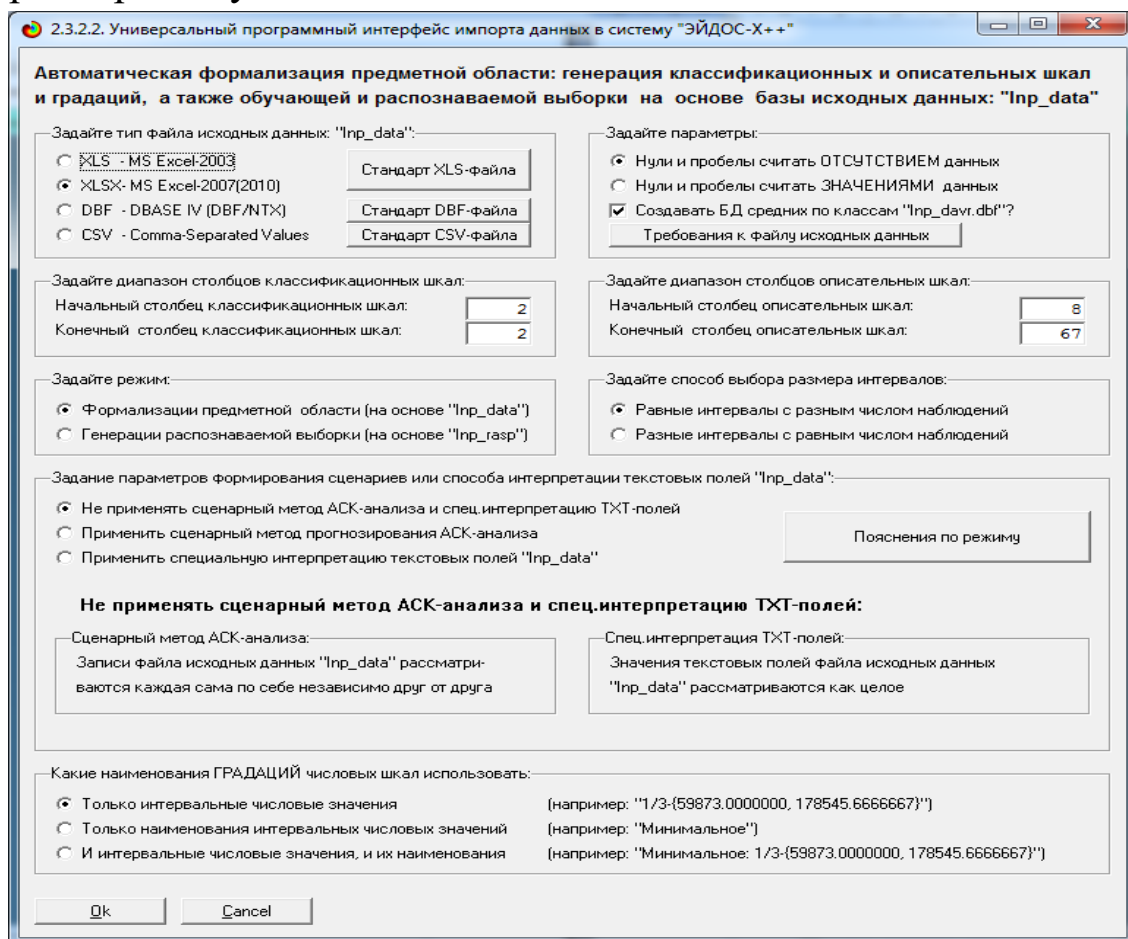


Рисунок 7. Главная экранная форма универсального программного интерфейса системы «Эйдос» с внешними базами данных (режим 2.3.2.2)

После запуска процесса ввода данных из файла: ..\AID_DATA\Inp_data\inp_data.xlsx в базы данных системы «Эйдос» определяется количество заданных текстовых и числовых классификационных и описательных шкал и градаций [7] и выводится окно внутреннего калькулятора данного режима, в котором мы можем задать число интервальных значений в числовых шкалах (рисунок 8):

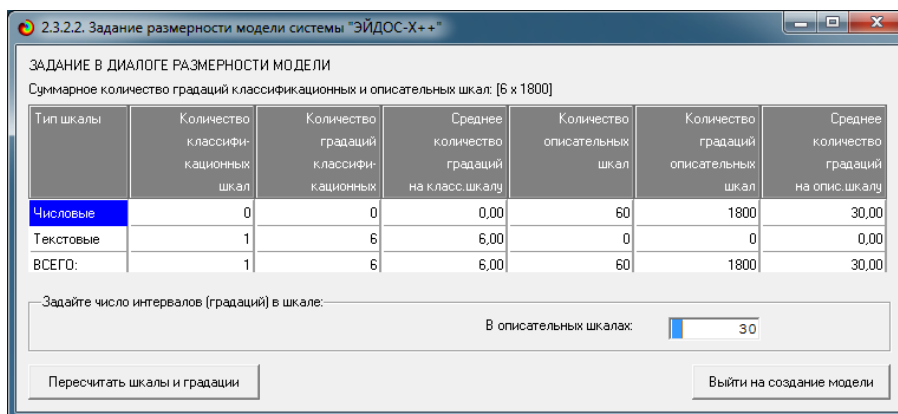


Рисунок 8. Экранная форма внутреннего калькулятора универсального программного интерфейса системы «Эйдос» с внешними базами данных

После клика по кнопке «Выйти на создание модели» начинается процесс импорта данных оцифровки изображений из файла `..\AID_DATA\Inp_data\inp_data.xlsx` в базы данных системы «Эйдос». При этом по сути происходит нормализация базы исходных данных, т.е. создаются справочники классификационных и описательных шкал и градаций и исходные данные кодируются с их использованием, в результате чего формируется обучающая выборка и база событий (эвентологическая база данных) (рисунок 9):

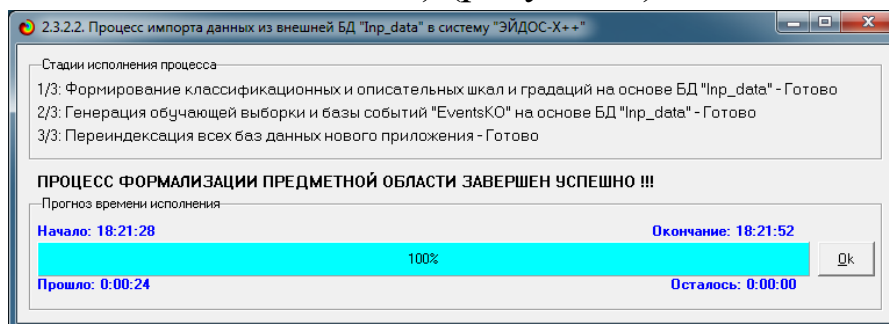


Рисунок 9. Экранная форма, отображающая этапы импорта данных из внешней базы данных в систему «Эйдос»

Сами справочники классификационных и описательных шкал и градаций, обучающая выборка и эвентологическая база данных могут быть просмотрены в режимах 2.1, 2.2, 2.3.1, 2.4 системы «Эйдос». В результате работы программного интерфейса с внешними базами данных 2.3.2.2 также формируется таблица

..\AID_DATA\Inp_data\inp_davr.xls с усредненными данными по классам.

Затем запускается режим 3.5, обеспечивающий синтез и верификацию (оценку достоверности) моделей (рисунок 10):

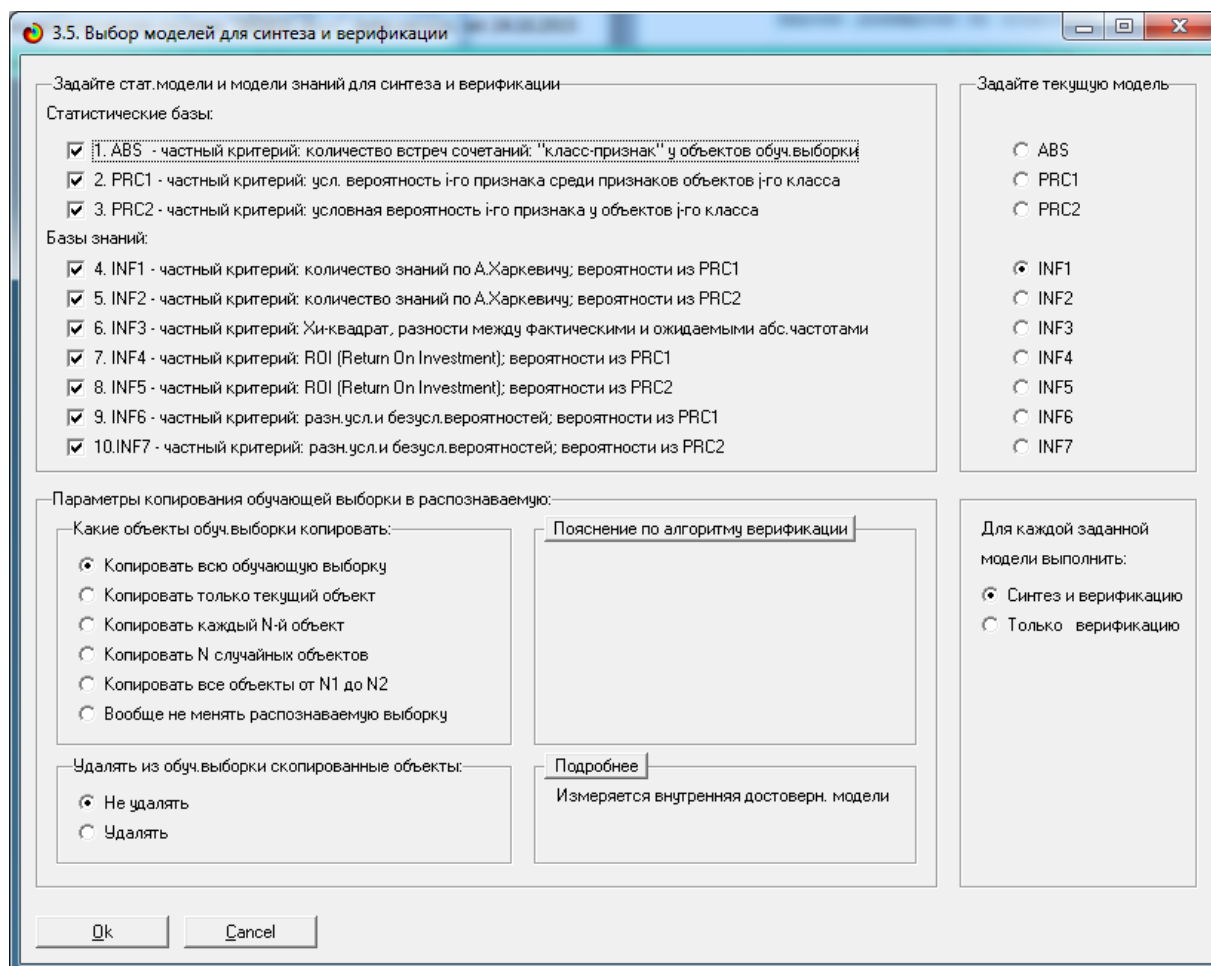


Рисунок 10. Экранная форма режима синтеза и верификации моделей системы «Эйдос» (режима 3.5)

Запускаем этот режим с параметрами по умолчанию.

В результате работы данного режима создаются и верифицируются три статистических модели (корреляционная матрица, матрицы условных и безусловных процентных распределений) и 7 системно-когнитивных моделей (моделей знаний) (рисунки 11 и 12) [7]:

Последовательность обработки данных, информации и знаний в системе «Эйдос-X++»

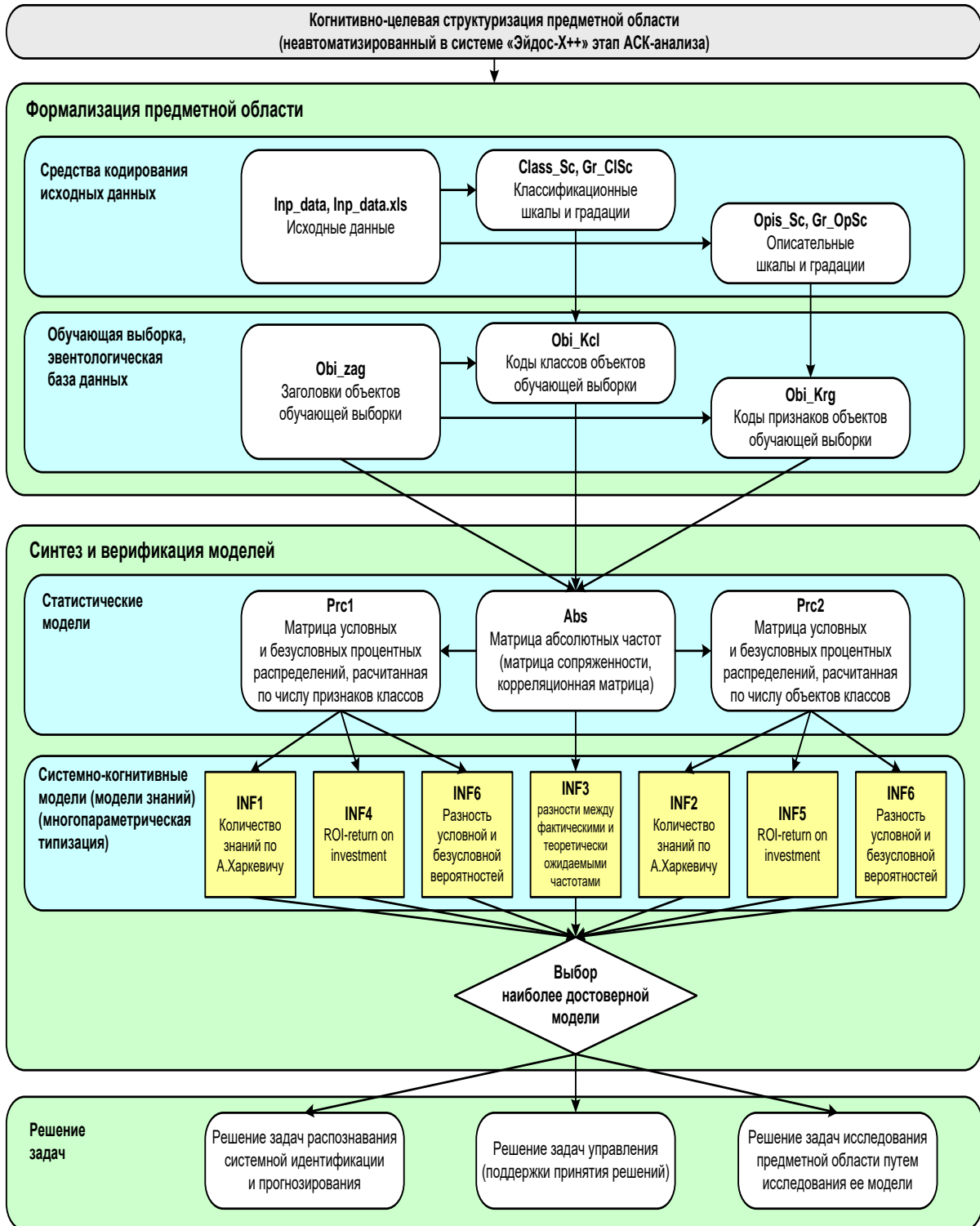


Рисунок 11. Этапы последовательного повышения степени формализации модели от данных к информации, а от нее к знаниям³

³ Подробнее об этом можно прочитать в работе [7] и других работах, посвященных АСК-анализу

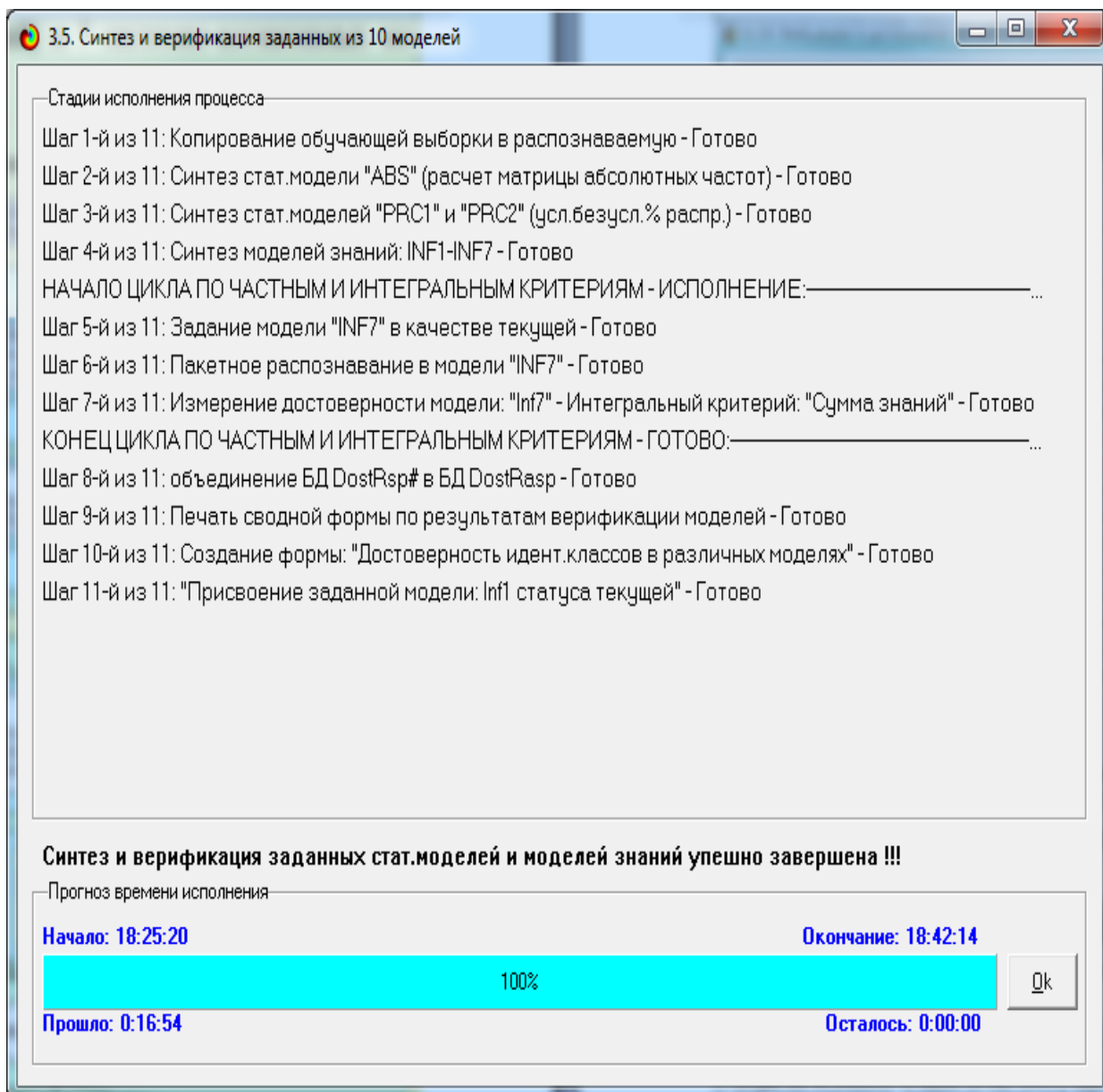


Рисунок 12. Экранная форма отображения стадии исполнения режима синтеза и верификации моделей

Из рисунка 12 видно, что процесс синтеза и верификации моделей на выборке из 451 листа занял 16 минут 54 секунды.

Достоверность моделей оценивается в этом же режиме 3.5 в соответствии с предложенной проф. Е.В. Луценко метрикой, сходной по смыслу с известным F-критерием, но не основанной на предположении о нормальности

распределения, независимости и аддитивности факторов (рисунки 13 и 14).

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентифика... | Вероятность правильной не идентиф... | Средняя вероятн... | Дата получения результата | Время получения результ... |
|---|-----------------------------------|--------------------------------------|--------------------------------------|--------------------|---------------------------|----------------------------|
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Корреляция абс. частот с обр... | 98.152 | 4.382 | 51.267 | 24.10.2015 | 18:30:36 |
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Сумма абс. частот по признак... | 100.000 | | 50.000 | 24.10.2015 | 18:30:36 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Корреляция усл.отн. частот с о... | 98.152 | 4.377 | 51.264 | 24.10.2015 | 18:31:55 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 24.10.2015 | 18:31:55 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Корреляция усл.отн. частот с о... | 98.152 | 4.382 | 51.267 | 24.10.2015 | 18:33:14 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 24.10.2015 | 18:33:14 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 66.359 | 90.918 | 78.638 | 24.10.2015 | 18:34:31 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 83.549 | 76.080 | 79.815 | 24.10.2015 | 18:34:31 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 66.359 | 90.918 | 78.638 | 24.10.2015 | 18:35:48 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 83.549 | 76.144 | 79.846 | 24.10.2015 | 18:35:48 |
| 6. INF3 - частный критерий: Хинкватрат, разности между фактич... | Семантический резонанс зна... | 88.355 | 64.566 | 76.460 | 24.10.2015 | 18:37:08 |
| 6. INF3 - частный критерий: Хинкватрат, разности между фактич... | Сумма знаний | 88.355 | 64.566 | 76.460 | 24.10.2015 | 18:37:08 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятн... | Семантический резонанс зна... | 49.538 | 97.155 | 73.346 | 24.10.2015 | 18:38:26 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятн... | Сумма знаний | 93.531 | 56.926 | 75.228 | 24.10.2015 | 18:38:26 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятн... | Семантический резонанс зна... | 49.538 | 97.155 | 73.346 | 24.10.2015 | 18:39:40 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятн... | Сумма знаний | 93.531 | 56.847 | 75.189 | 24.10.2015 | 18:39:40 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Семантический резонанс зна... | 85.582 | 66.584 | 76.083 | 24.10.2015 | 18:40:57 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Сумма знаний | 88.355 | 59.010 | 73.682 | 24.10.2015 | 18:40:58 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Семантический резонанс зна... | 85.582 | 66.647 | 76.115 | 24.10.2015 | 18:42:14 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Сумма знаний | 88.355 | 59.010 | 73.682 | 24.10.2015 | 18:42:14 |

Рисунок 13. Оценка достоверности моделей с помощью непараметрической метрики, сходной с F-критерием

Режим: Помощь по режиму: 4.1.3.6: Виды прогнозов и принцип определения достоверности моделей в системе "Эйдос-Х++".

ПОЛОЖИТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Предположим, модель дает такой прогноз: выпадет 1, 2, 3, 4, 5 или 6. В этом случае у нее будет 100% достоверность идентификации, т.е. не будет ни одного объекта, не отнесенного к тому классу, к которому он действительно относится, но при этом будет очень большая ошибка ложной идентификации, т.к. огромное количество объектов будет отнесено к классам, к которым они не относятся (и именно за счет этого у модели и будет очень высокая достоверность идентификации). Ясно, что такой прогноз бесполезен, поэтому он и назван мной псевдопрогнозом.

ОТРИЦАТЕЛЬНЫЙ ПСЕВДОПРОГНОЗ.
Представим себе, что мы выбрасываем кубик с 6 гранями, и модель предсказывает, что не выпадет: 1, 2, 3, 4, 5 и 6, а что-то из этого естественно выпало. Конечно, модель дает ошибку в прогнозе в том плане, что не предсказала, что выпадет, зато она очень хорошо угадала, что не выпадет. Но ясно, что выпадет что-то одно, а не все, что предсказано, поэтому такого рода предсказания хорошо оправдываются в том, что не произошло и плохо в том, что произошло, т.е. в этом случае у модели будет 100% достоверность не идентификации, но очень низкая достоверность идентификации.

ИДЕАЛЬНЫЙ ПРОГНОЗ.
Если в случае с кубиком мы прогнозируем, что выпадет, например 1, и соответственно прогнозируем, что не выпадет 2, 3, 4, 5, и 6, то это идеальный прогноз, имеющий, если он осуществляется, 100% достоверность идентификации и не идентификации. Идеальный прогноз, который полностью снимает неопределенность о будущем состоянии объекта прогнозирования, на практике удается получить крайне редко и обычно мы имеем дело с реальным прогнозом.

РЕАЛЬНЫЙ ПРОГНОЗ.
На практике мы чаще всего сталкиваемся именно с этим видом прогноза. Реальный прогноз уменьшает неопределенность о будущем состоянии объекта прогнозирования, но не полностью, как идеальный прогноз, а оставляет некоторую неопределенность не снятой. Например, для игрального кубика делается такой прогноз: выпадет 1 или 2, и, соответственно, не выпадет 3, 4, 5 или 6. Понятно, что полностью на практике такой прогноз не может осуществиться, т.к. варианты выпадения кубика альтернативны, т.е. не может выпасть одновременно и 1, и 2. Поэтому у реального прогноза всегда будет определенная ошибка идентификации. Соответственно, если не осуществится один или несколько из прогнозируемых вариантов, то возникнет и ошибка не идентификации, т.к. это не прогнозировалось моделью. Теперь представьте себе, что у вас не 1 кубик и прогноз его поведения, а тысячи. Тогда можно посчитать средневзвешенные характеристики всех этих видов прогнозов.

Таким образом, если просуммировать проценты верной идентификации и не идентификации и вычесть проценты ложной идентификации и ложной не идентификации, то это и будет критерий качества модели, учитывающий как ее способность верно относить объекты к классам, которым они относятся, так и ее способность верно не относить объекты к тем классам, к которым они не относятся. Ясно, что этот критерий очень сходен по смыслу с известным F-критерием и сходные оценки качества моделей.

Рисунок 14. Help режима 4.1.3.6: пояснение смысла непараметрической метрики, сходной с F-критерием

Из рисунка 13 мы видим, что наиболее достоверная модель, основанная на модифицированной мере Харкевича [7], дает среднюю достоверность определения сорта по листу около 80%, причем достоверность правильного отнесения листа к сорту, к которому он относится, составляет 83%, а правильного не отнесения к сорту, к которому он не относится – около 79%.

В соответствии со схемой, приведенной на рисунке 11, и информацией по достоверности моделей, приведенной на рисунке 13, в режиме 5.6 системы «Эйдос» зададим системно-когнитивную модель INF2 в качестве текущей (рисунок 15) и проведем в ней пакетную идентификацию в режиме 4.1.2 (рисунок 16):

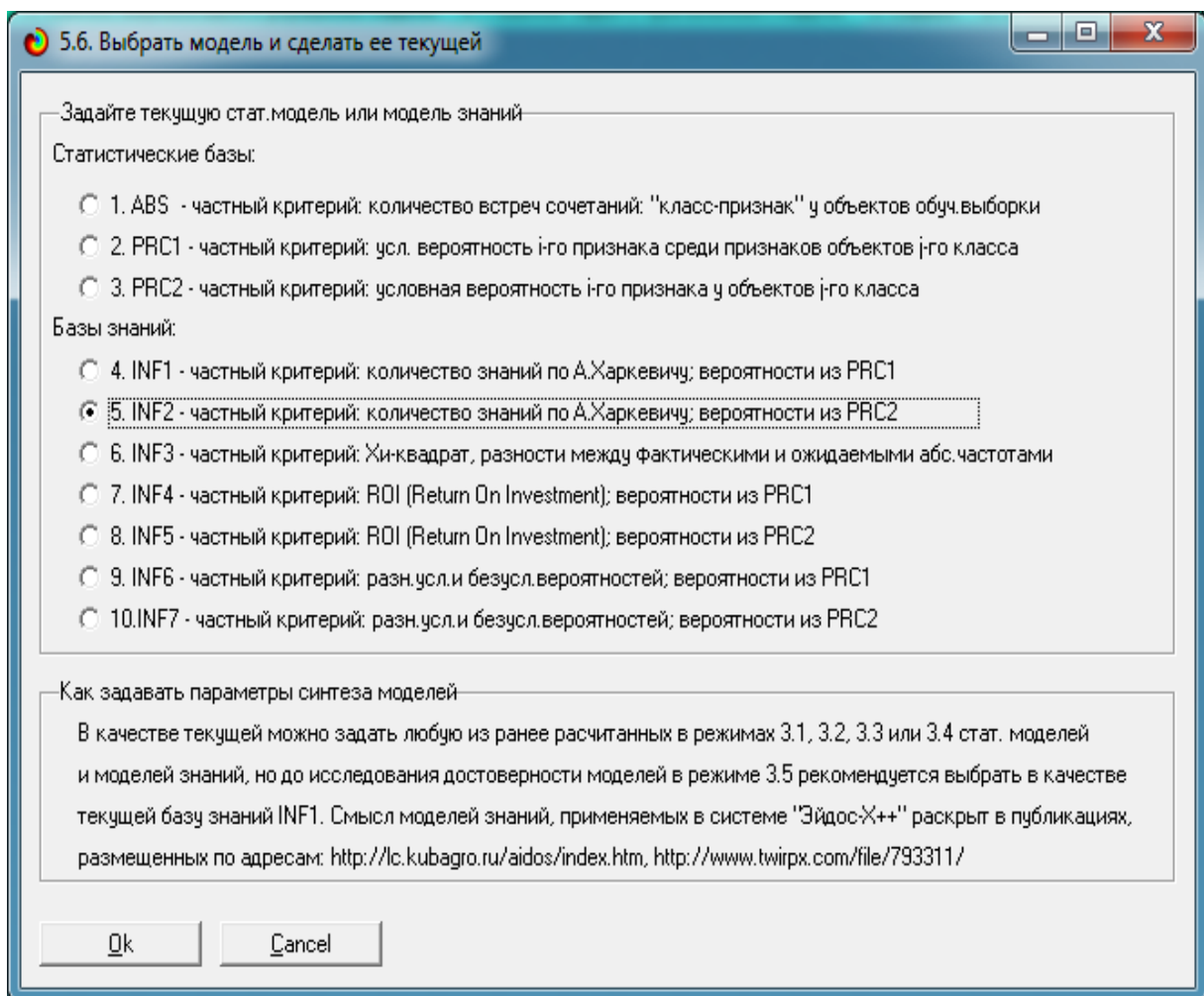


Рисунок 15. Экранная форма, позволяющая задать любую модель в качестве текущей

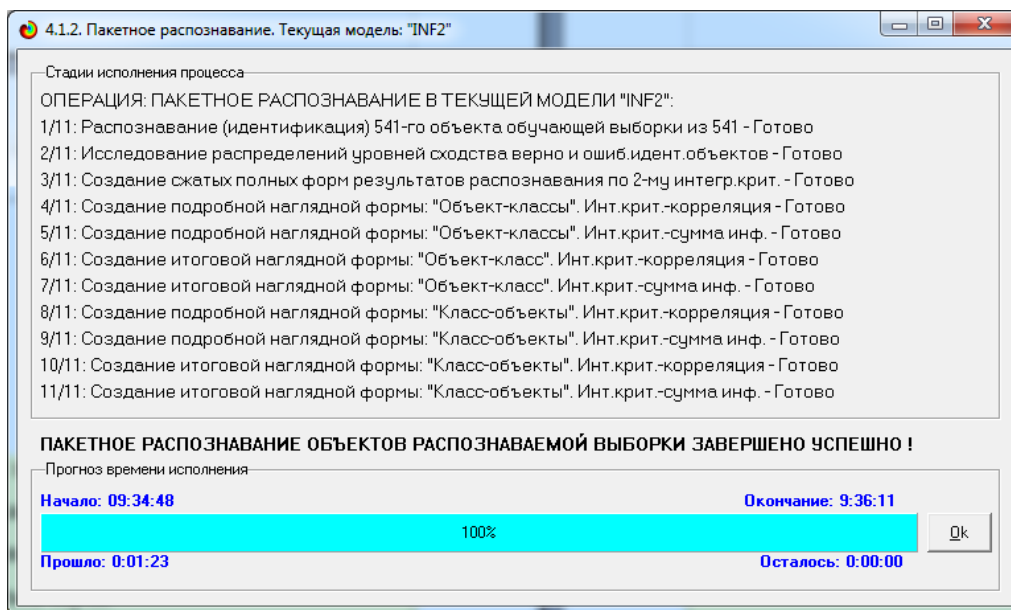


Рисунок 16. Экранная форма с отображением этапов и стадии решения задачи идентификации листьев с обобщенными образами листьев сортов

Из данной экранной формы видно, что идентификация 541 листа с обобщенными образами листьев сортов выполнена за минуту 23 секунды.

На рисунке 17 представлен пример результатов идентификации:

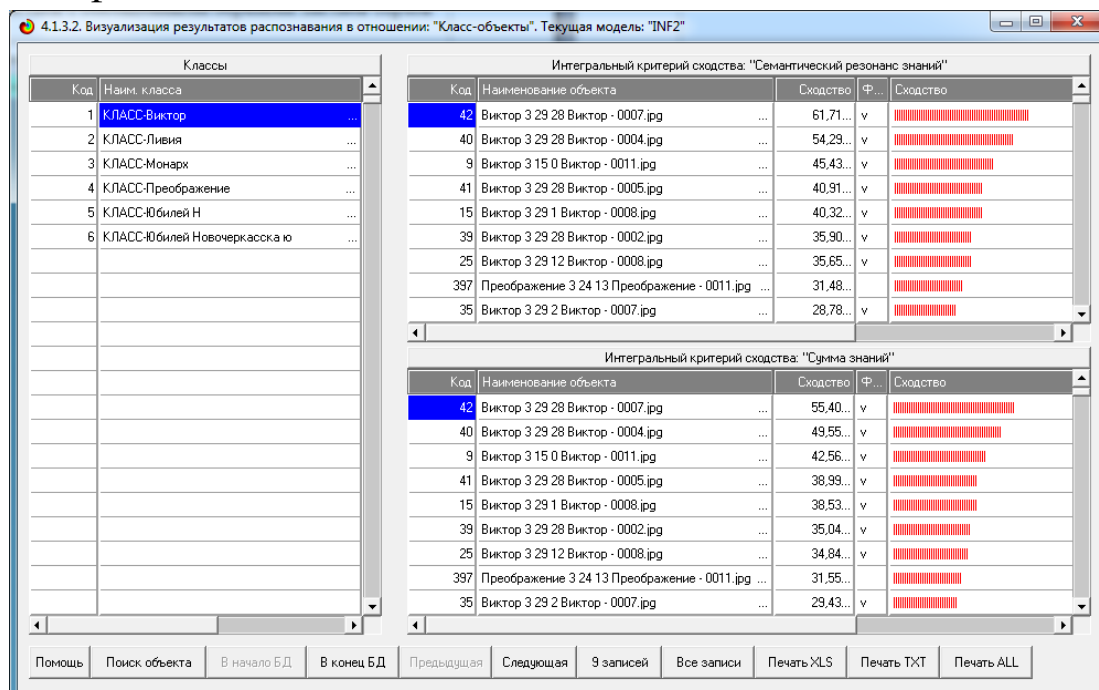


Рисунок 17. Экранная форма с результатами идентификации листьев с обобщенным образом листьев сорта «Виктор»

При идентификации возникают ошибки неидентификации и ложной идентификации, снижающие достоверность модели. По мнению авторов, основной причиной этих ошибок является наличие в обучающей выборке по каждому сорту винограда «нетипичных» для данного сорта листьев, значительно отличающихся по своей форме от основной массы листьев сорта (явление полиморфизма). По сути это связано с не очень высоким качеством обучающей выборки.

В системе «Эйдос» есть много различных средств повышения качества моделей, в частности, за счет улучшения качества обучающей выборки. Для поиска и удаления из обучающей выборки данных по нетипичным листьям может быть применен режим: 3.7.6. Разделение классов на типичную и нетипичную части. Этот режим был реализован в DOS-версии системы «Эйдос» [8, 9, 10], а во время работы над статьей он был реализован в несколько усовершенствованном виде и в новой версии системы «Эйдос-X++» [11]. На рисунке 18 приведена экранная форма этого режима:

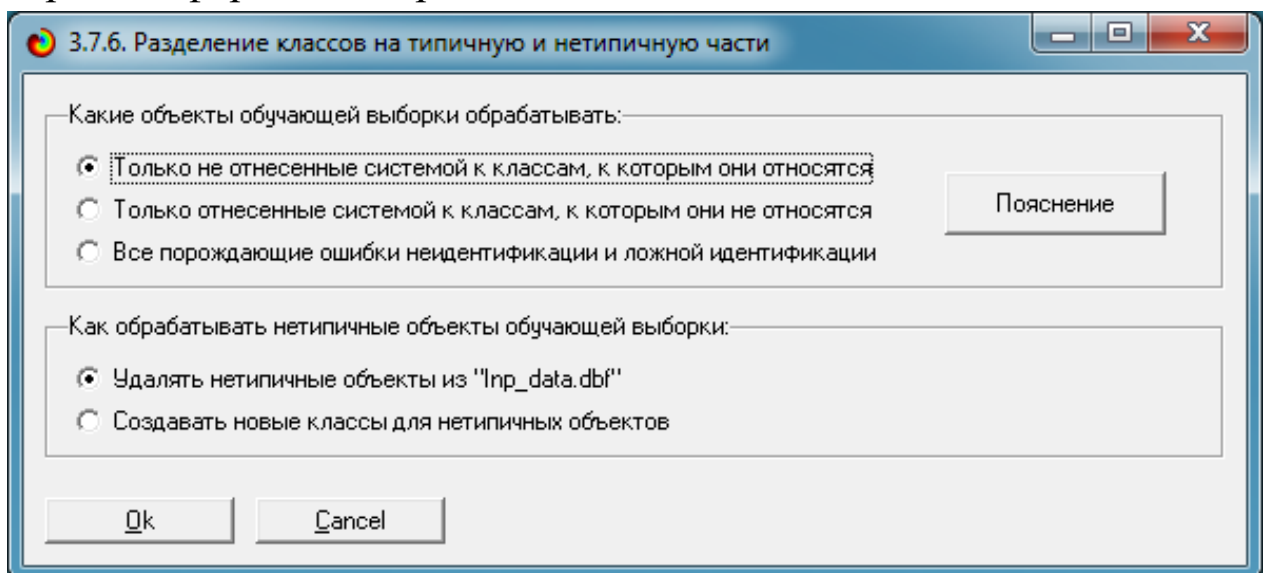


Рисунок 18. Экранная форма режима: 3.7.6. Разделение классов на типичную и нетипичную части

Вместо описания данного режима приведем его Help (рисунок 19):

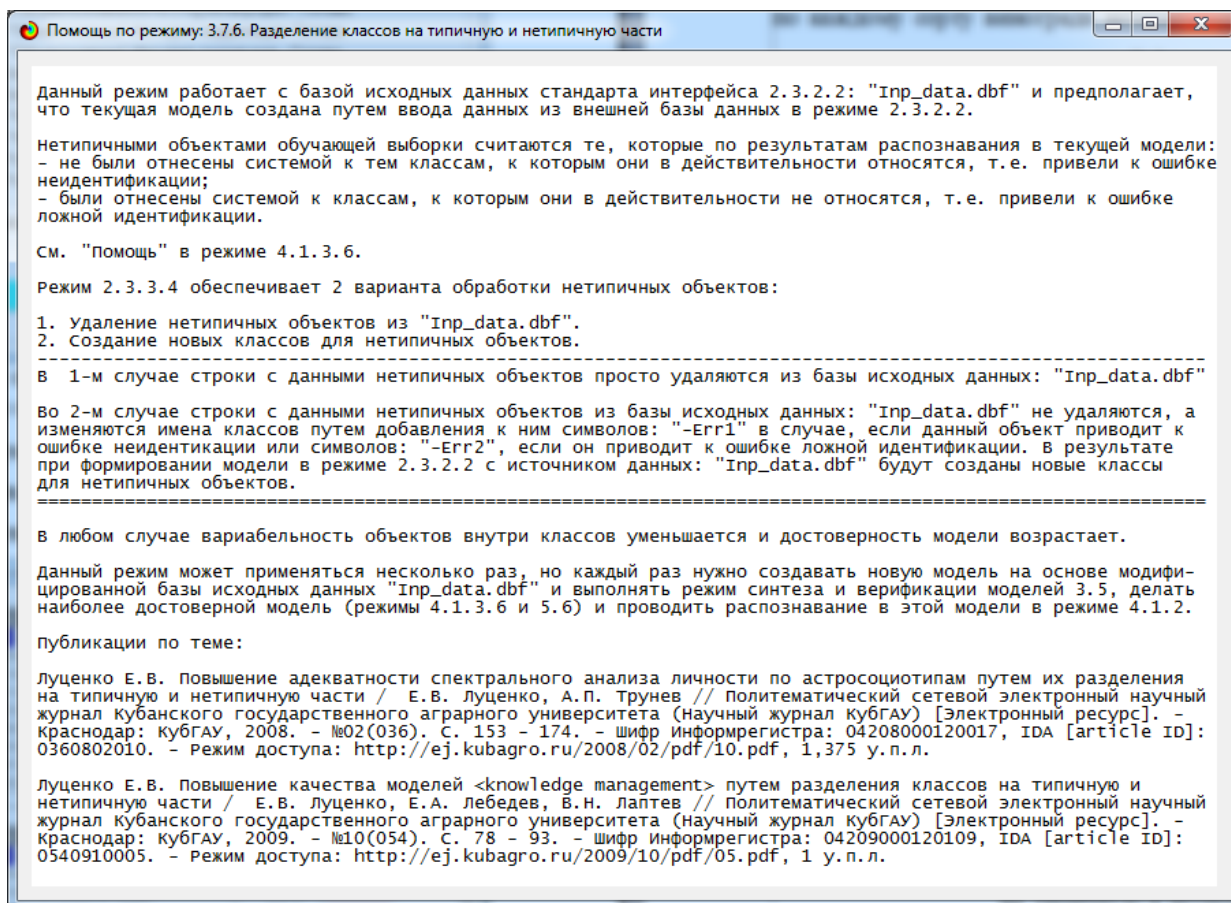


Рисунок 19. Help режима: 3.7.6. Разделение классов на типичную и нетипичную части

В результате работы данного режима при параметрах, приведенных в экранной форме на рисунке 18, из обучающей выборки было убрано 182 нетипичных изображения листьев из 541, в результате чего осталось 359 изображений (рисунок 20):

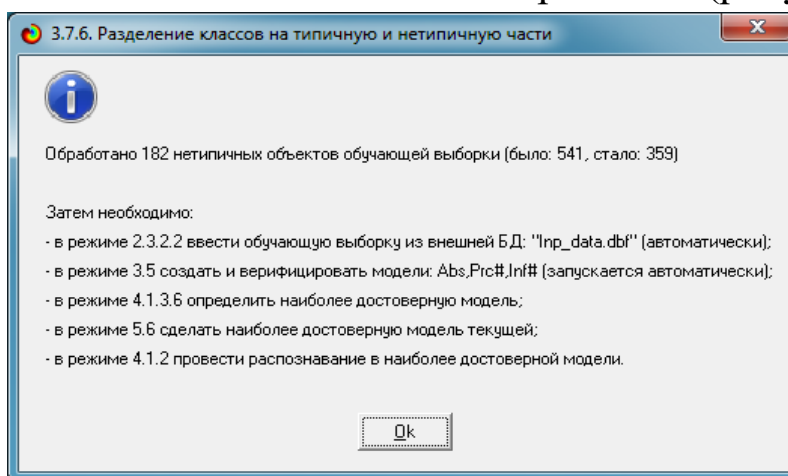


Рисунок 20. Экранная форма с результатами удаления нетипичных объектов обучающей выборки и рекомендациями по последующим действиям

улучшенной обучающей выборки создается новое приложение, а потом это новое приложение готовится для следующего применения режима 3.7.6 (см. ниже рисунок 20).

4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит.. Текущая модель: "INF2"

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентифика... | Вероятность правильной не идентиф... | Средняя вероятно... результ... | Дата получения результата | Время получения результ... |
|---|-----------------------------------|--------------------------------------|--------------------------------------|--------------------------------|---------------------------|----------------------------|
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Корреляция абс. частот с обр... | 98.607 | 11.347 | 54.977 | 25.10.2015 | 10:18:26 |
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Сумма абс. частот по признак... | 100.000 | | 50.000 | 25.10.2015 | 10:18:26 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Корреляция усл.отн. частот с о... | 98.607 | 11.122 | 54.865 | 25.10.2015 | 10:19:17 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 25.10.2015 | 10:19:17 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Корреляция усл.отн. частот с о... | 98.607 | 11.347 | 54.977 | 25.10.2015 | 10:20:05 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 25.10.2015 | 10:20:05 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 89.136 | 77.272 | 83.204 | 25.10.2015 | 10:20:52 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 92.479 | 67.841 | 80.160 | 25.10.2015 | 10:20:52 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 89.136 | 77.272 | 83.204 | 25.10.2015 | 10:21:41 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 92.479 | 67.841 | 80.160 | 25.10.2015 | 10:21:41 |
| 6. INF3 - частный критерий: Хи-квадрат, разности между фактич... | Семантический резонанс зна... | 96.379 | 63.072 | 79.725 | 25.10.2015 | 10:22:29 |
| 6. INF3 - частный критерий: Хи-квадрат, разности между фактич... | Сумма знаний | 96.379 | 63.072 | 79.725 | 25.10.2015 | 10:22:29 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 86.072 | 82.777 | 84.425 | 25.10.2015 | 10:23:17 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Сумма знаний | 99.443 | 51.485 | 75.464 | 25.10.2015 | 10:23:17 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 86.072 | 82.777 | 84.425 | 25.10.2015 | 10:24:06 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятно... | Сумма знаний | 99.443 | 51.485 | 75.464 | 25.10.2015 | 10:24:06 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Семантический резонанс зна... | 94.708 | 62.113 | 78.410 | 25.10.2015 | 10:24:55 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Сумма знаний | 96.379 | 54.455 | 75.417 | 25.10.2015 | 10:24:55 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Семантический резонанс зна... | 94.708 | 62.113 | 78.410 | 25.10.2015 | 10:25:43 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Сумма знаний | 96.379 | 54.455 | 75.417 | 25.10.2015 | 10:25:43 |

Помощь

4.1.3.6. Обобщ.форма по достов.моделей при разн.инт.крит.. Текущая модель: "INF4"

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентифика... | Вероятность правильной не идентиф... | Средняя вероятно... результ... | Дата получения результата | Время получения результ... |
|---|-----------------------------------|--------------------------------------|--------------------------------------|--------------------------------|---------------------------|----------------------------|
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Корреляция абс. частот с обр... | 98.706 | 16.636 | 57.671 | 25.10.2015 | 10:53:25 |
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Сумма абс. частот по признак... | 100.000 | | 50.000 | 25.10.2015 | 10:53:25 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Корреляция усл.отн. частот с о... | 99.029 | 16.621 | 57.825 | 25.10.2015 | 10:54:08 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 25.10.2015 | 10:54:08 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Корреляция усл.отн. частот с о... | 98.706 | 16.636 | 57.671 | 25.10.2015 | 10:54:50 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Сумма усл.отн. частот по приз... | 100.000 | | 50.000 | 25.10.2015 | 10:54:50 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 93.204 | 76.086 | 84.645 | 25.10.2015 | 10:55:32 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 97.087 | 66.757 | 81.922 | 25.10.2015 | 10:55:32 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 93.204 | 76.086 | 84.645 | 25.10.2015 | 10:56:16 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 97.087 | 66.757 | 81.922 | 25.10.2015 | 10:56:16 |
| 6. INF3 - частный критерий: Хи-квадрат, разности между фактич... | Семантический резонанс зна... | 99.676 | 64.870 | 82.273 | 25.10.2015 | 10:56:58 |
| 6. INF3 - частный критерий: Хи-квадрат, разности между фактич... | Сумма знаний | 99.676 | 64.870 | 82.273 | 25.10.2015 | 10:56:58 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 95.793 | 80.886 | 88.339 | 25.10.2015 | 10:57:40 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Сумма знаний | 100.000 | 51.355 | 75.677 | 25.10.2015 | 10:57:40 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 95.793 | 80.886 | 88.339 | 25.10.2015 | 10:58:22 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятно... | Сумма знаний | 100.000 | 51.301 | 75.650 | 25.10.2015 | 10:58:22 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Семантический резонанс зна... | 98.382 | 63.615 | 80.998 | 25.10.2015 | 10:59:04 |
| 9. INF6 - частный критерий: разн. усл. и безуслов. вероятностей; вер... | Сумма знаний | 99.676 | 53.600 | 76.638 | 25.10.2015 | 10:59:04 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Семантический резонанс зна... | 98.382 | 63.615 | 80.998 | 25.10.2015 | 10:59:46 |
| 10. INF7 - частный критерий: разн. усл. и безуслов. вероятностей; ве... | Сумма знаний | 99.676 | 53.600 | 76.638 | 25.10.2015 | 10:59:46 |

Помощь

Рисунок 21. Оценка достоверности моделей, созданных на основе обучающей выборки с удаленными данными о нетипичных листьях на 1-й и 2-й итерациях (непараметрическая метрика, сходная с F-критерием)

Из рисунка 21 видно, что в наилучшей модели, созданной на 2-й итерации (INF4), достоверность идентификации, определения сорта винограда по форме листа (контуру) составляет 95,8%, а достоверность неидентификации, т.е. определения сортов, к которым данный лист не относится – 80,9%. Это уже результат, достаточно высокий для того, чтобы можно было говорить не только о научной новизне, но и о практической значимости предлагаемых подходов к решению некоторых сформулированных в работе важных задач ампелографии.

Отметим также, что дальнейшие итерации в данном случае проводить нецелесообразно, т.к. наиболее эффективны, т.е. наиболее сильно повышают достоверность моделей именно первые итерации, а последующие все в меньшей и меньшей степени, что вполне естественно, ведь достоверность модели асимптотически приближается к 100%. Кроме того, в результате проведенных итераций уже достигнут достаточный уровень достоверности. В исходной обучающей выборке был 541 объект (лист), в модели 1-й итерации – 359, а в модели 2-й итерации – 309.

Выберем в качестве текущей модель INF4 2-й итерации и решим в ней сформулированные в начале статьи важные задачи ампелографии:

- 1) оцифровка сканированных изображений листьев и создание их математических моделей;
- 2) формирование математических моделей конкретных листьев с применением теории информации;
- 3) формирование моделей обобщенных образов листьев различных сортов;
- 4) сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу;

5) количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов.

Задача 1: оцифровка сканированных изображений листьев и создание их математических моделей.

Решение этой задачи подробно описано в начале статьи. Оцифровка сканированных изображений осуществляется модулем 2.3.2.4, программная реализация которого разработана Д.К. Бандык по постановке и алгоритму проф. Е.В. Луценко [3, 4]. В результате работы модуля создается Excel-файл обучающей выборки (таблица 1), в строках которого содержатся следующие количественные данные об изображениях, представляющие их математическую модель (при заданных параметрах оцифровки):

Колонки с 1-й по 7:

1. Объект обучающей выборки
2. Класс
3. X центра тяжести
4. Y центра тяжести
5. Площадь (пикс.)
6. Среднее
7. Ср. кв. откл.

Колонки с 8-й по 67:

| | | | | | | | | | |
|-----|------|-----|------|-----|------|-----|------|-----|------|
| 8. | 000° | 9. | 006° | 10. | 012° | 11. | 018° | 12. | 024° |
| 13. | 030° | 14. | 036° | 15. | 042° | 16. | 048° | 17. | 054° |
| 18. | 060° | 19. | 066° | 20. | 072° | 21. | 078° | 22. | 084° |
| 23. | 090° | 24. | 096° | 25. | 102° | 26. | 108° | 27. | 114° |
| 28. | 120° | 29. | 126° | 30. | 132° | 31. | 138° | 32. | 144° |
| 33. | 150° | 34. | 156° | 35. | 162° | 36. | 168° | 37. | 174° |
| 38. | 180° | 39. | 186° | 40. | 192° | 41. | 198° | 42. | 204° |
| 43. | 210° | 44. | 216° | 45. | 222° | 46. | 228° | 47. | 234° |
| 48. | 240° | 49. | 246° | 50. | 252° | 51. | 258° | 52. | 264° |
| 53. | 270° | 54. | 276° | 55. | 282° | 56. | 288° | 57. | 294° |
| 58. | 300° | 59. | 306° | 60. | 312° | 61. | 318° | 62. | 324° |
| 63. | 330° | 64. | 336° | 65. | 342° | 66. | 348° | 67. | 354° |

Задача 2: формирование математических моделей конкретных листьев с применением теории информации.

В результате выполнения режима 3.5. Синтез и верификация моделей, создано 3 статистических модели и 7 когнитивных моделей исследуемой предметной области. Наименования этих моделей приведены на рисунках 10 и 11, а подробнее они описаны в работе [7] и других работах, посвященных АСК-анализу⁵. В таблице 3 приведен фрагмент базы знаний модели INF1 на 2-й итерации.

Колонки приведенной базы знаний соответствуют классам, в данной работе это сорта винограда.

Строки приведенной базы знаний представляют собой градации описательных шкал, в данной работе это интервальные значения расстояний от центра тяжести листа до его контура в пикселях при различных углах поворота радиуса-вектора.

Значения на пересечениях строк и колонок в модели INF1 представляют собой количество информации по А.Харкевичу [7] в элементе контура о принадлежности листа с данным элементом контура к сорту винограда. Это количество информации выражено в миллибитах, т.е. тысячных долях бита с целью экономии знакомест в таблице (т.к. все значения начинаются с нуля и запятой). *Знак и модуль значения отражают степень характерности или не характерности соответствующего элемента контура данному сорту винограда.* Положительные значения говорят о том, что данный элемент контура характерен для данного сорта винограда, а отрицательные – о том, что он не характерен. Степень характерности или не характерности выражается модулем значения. Пустые значения соответствуют случаю, когда данный элемент контура листа никогда не встречался у данного сорта по данным обучающей выборки. Положительные значения получаются, если данный элемент контура встречается у данного

⁵ Ссылки на многие из них есть на сайте: <http://lc.kubagro.ru/>

сорта с более высокой вероятностью, чем в среднем по выборке, а отрицательные – если реже, чем в среднем по выборке. Подробнее о способах расчета статистических и когнитивных моделей в АСК-анализе и системе «Эйдос» можно ознакомиться в работе [7] и других работах по этой проблематике.

Таблица 3. - Модель INF1 на 2-й итерации (фрагмент) (миллибиты)

| Код | Описательные шкалы и градации | Классы (сорта винограда) | | | | | |
|-----|--------------------------------------|--------------------------|-------|--------|--------------|----------|-----------------------|
| | | Виктор | Ливия | Монарх | Преображение | Юбилей Н | Юбилей Новочаркаска ю |
| 1 | 000*1/30-{233.1520386, 249.7032003} | 89 | | | | 336 | |
| 2 | 000*2/30-{249.7032003, 266.2543620} | 148 | | | | 212 | 490 |
| 3 | 000*3/30-{266.2543620, 282.8055237} | 148 | 390 | | | -153 | |
| 4 | 000*4/30-{282.8055237, 299.3566854} | 148 | 196 | | | 136 | |
| 5 | 000*5/30-{299.3566854, 315.9078471} | 165 | 165 | -452 | | 47 | 102 |
| 6 | 000*6/30-{315.9078471, 332.4590068} | 205 | -67 | -443 | | 162 | 110 |
| 7 | 000*7/30-{332.4590068, 349.0101705} | 213 | -28 | -404 | | 178 | |
| 8 | 000*8/30-{349.0101705, 365.5613322} | 161 | -320 | -91 | | 167 | |
| 9 | 000*9/30-{365.5613322, 382.1124939} | -80 | | 91 | | 132 | |
| 10 | 000*10/30-{382.1124939, 399.6636566} | -208 | -440 | 31 | | 97 | -49 |
| 11 | 000*11/30-{399.6636566, 415.2148173} | | | 225 | | 259 | 18 |
| 12 | 000*12/30-{415.2148173, 431.7659790} | -141 | -382 | 136 | | -441 | 383 |
| 13 | 000*13/30-{431.7659790, 448.3171407} | -130 | -82 | 110 | | -430 | 30 |
| 14 | 000*14/30-{448.3171407, 464.8683024} | -90 | 191 | -3 | | -380 | 80 |
| 15 | 000*15/30-{464.8683024, 481.4194641} | -248 | 117 | 78 | | 383 | |
| 16 | 000*16/30-{481.4194641, 497.9706258} | -93 | 213 | 19 | | | 208 |
| 17 | 000*17/30-{497.9706258, 514.5217875} | | 183 | 113 | | | 120 |
| 18 | 000*18/30-{514.5217875, 531.0729492} | | 454 | | | | |
| 19 | 000*19/30-{531.0729492, 547.6241109} | | | 153 | | 105 | |
| 20 | 000*20/30-{547.6241109, 564.1752726} | | 320 | 19 | | | |
| 21 | 000*21/30-{564.1752726, 580.7264343} | | | | | | |
| 22 | 000*22/30-{580.7264343, 597.2775960} | | | 260 | | | |
| 23 | 000*23/30-{597.2775960, 613.8287577} | | | | | | |
| 24 | 000*24/30-{613.8287577, 630.3799194} | | | | | | |
| 25 | 000*25/30-{630.3799194, 646.9310811} | | | | | | |
| 26 | 000*26/30-{646.9310811, 663.4822428} | | | | | | |
| 27 | 000*27/30-{663.4822428, 680.0334045} | | | | | | |
| 28 | 000*28/30-{680.0334045, 696.5845662} | | | | | | |
| 29 | 000*29/30-{696.5845662, 713.1357279} | | | | | | |
| 30 | 000*30/30-{713.1357279, 729.6868896} | | | | | 394 | |
| 31 | 006*1/30-{208.7764567, 224.0454111} | | | 165 | | 288 | |
| 32 | 006*2/30-{224.0454111, 239.3143636} | 513 | | | | | |
| 33 | 006*3/30-{239.3143636, 254.5833160} | | | 347 | | 105 | |
| 34 | 006*4/30-{254.5833160, 269.8522684} | 183 | | | | 172 | 484 |
| 35 | 006*5/30-{269.8522684, 285.1212209} | 330 | -17 | -211 | | -77 | |
| 36 | 006*6/30-{285.1212209, 300.3901733} | 183 | 183 | | | 124 | |
| 37 | 006*7/30-{300.3901733, 315.6591257} | 196 | 31 | | | 212 | |
| 38 | 006*8/30-{315.6591257, 330.9280782} | 1 | 272 | -252 | | -11 | |
| 39 | 006*9/30-{330.9280782, 346.1970306} | 41 | -124 | -136 | | 181 | 54 |
| 40 | 006*10/30-{346.1970306, 361.4659830} | -288 | -347 | -70 | | 224 | 195 |
| 41 | 006*11/30-{361.4659830, 376.7349355} | -217 | -275 | 42 | | 30 | 307 |
| 42 | 006*12/30-{376.7349355, 392.0038879} | 41 | 19 | 78 | | 249 | 81 |
| 43 | 006*13/30-{392.0038879, 407.2728403} | 41 | -288 | 95 | | -77 | 102 |
| 44 | 006*14/30-{407.2728403, 422.5417928} | -17 | | 178 | | -318 | 142 |
| 45 | 006*15/30-{422.5417928, 437.8107452} | -23 | -82 | 110 | | -141 | 30 |
| 46 | 006*16/30-{437.8107452, 453.0796976} | | -124 | 153 | | 383 | 54 |
| 47 | 006*17/30-{453.0796976, 468.3486501} | 1 | 18 | 80 | | 54 | 135 |
| 48 | 006*18/30-{468.3486501, 483.6176025} | -93 | 272 | 19 | | | |
| 49 | 006*19/30-{483.6176025, 498.8865549} | -141 | 224 | -28 | | 201 | 160 |
| 50 | 006*20/30-{498.8865549, 514.1555074} | -85 | 347 | -318 | | | 236 |
| 51 | 006*21/30-{514.1555074, 529.4244598} | | 165 | 105 | | -184 | |
| 52 | 006*22/30-{529.4244598, 544.6934122} | | | -17 | 212 | | |
| 53 | 006*23/30-{544.6934122, 559.9623647} | | | | | | |
| 54 | 006*24/30-{559.9623647, 575.2313171} | | | | | | |
| 55 | 006*25/30-{575.2313171, 590.5002695} | | | | | | |
| 56 | 006*26/30-{590.5002695, 605.7692220} | | | 78 | | 212 | |
| 57 | 006*27/30-{605.7692220, 621.0381744} | | | | | | |
| 58 | 006*28/30-{621.0381744, 636.3071268} | | | | | | |
| 59 | 006*29/30-{636.3071268, 651.5760793} | | | | | | |
| 60 | 006*30/30-{651.5760793, 666.8450317} | | | | | 394 | |
| 61 | 012*1/30-{116.6740494, 137.0900251} | | 454 | | | | |
| 62 | 012*2/30-{137.0900251, 157.5060008} | | | | | | |
| 63 | 012*3/30-{157.5060008, 177.9219765} | | | | | | |
| 64 | 012*4/30-{177.9219765, 198.3379522} | | | | | | |
| 65 | 012*5/30-{198.3379522, 218.7539279} | | | | | | 814 |
| 66 | 012*6/30-{218.7539279, 239.1699036} | 513 | | | | | |
| 67 | 012*7/30-{239.1699036, 259.5858793} | 224 | | | | 288 | |
| 68 | 012*8/30-{259.5858793, 280.0018550} | 272 | 320 | | | | |
| 69 | 012*9/30-{280.0018550, 300.4178307} | 107 | 231 | | | 124 | |

Всего в этой базе знаний 1800 записей, т.е. в таблице 3 приведено 3,8% ее объема.

Модели конкретных листьев с применением теории информации представляют собой количество информации в элементах их контуров о том, что эти листья принадлежат определенным сортам. В соответствии с подходом, развиваемым в АСК-анализе и реализованным в системе «Эйдос», считается, что лист относится к тому сорту винограда, о принадлежности к которому в элементах его контура содержится больше всего информации, т.е. используется аддитивный *интегральный критерий* [7].

Задача 3: формирование моделей обобщенных образов листьев различных сортов.

Модели обобщенных образов листьев содержатся в когнитивных моделях INF1 – INF7, фрагмент одной из которых приведен в таблице 3. Такая модель представляет собой вектор, координатами которого являются *частные критерии*, в частности в моделях INF1 и INF2 – количество информации [7]. Синтез и верификация этих моделей осуществляется в режиме 3.5 и подробно описаны в данной работе выше. Наименования этих моделей приведены на рисунках 10 и 11, а подробнее они описаны в работе [7] и других работах, посвященных АСК-анализу⁶.

В системе «Эйдос» реализован режим 4.7. АСК-анализ изображений, в котором можно выполнить практически все операции по оцифровке и анализу изображений, описанные в данной статье (рисунок 22):

⁶ Ссылки на многие из них есть на сайте: <http://lc.kubagro.ru/>

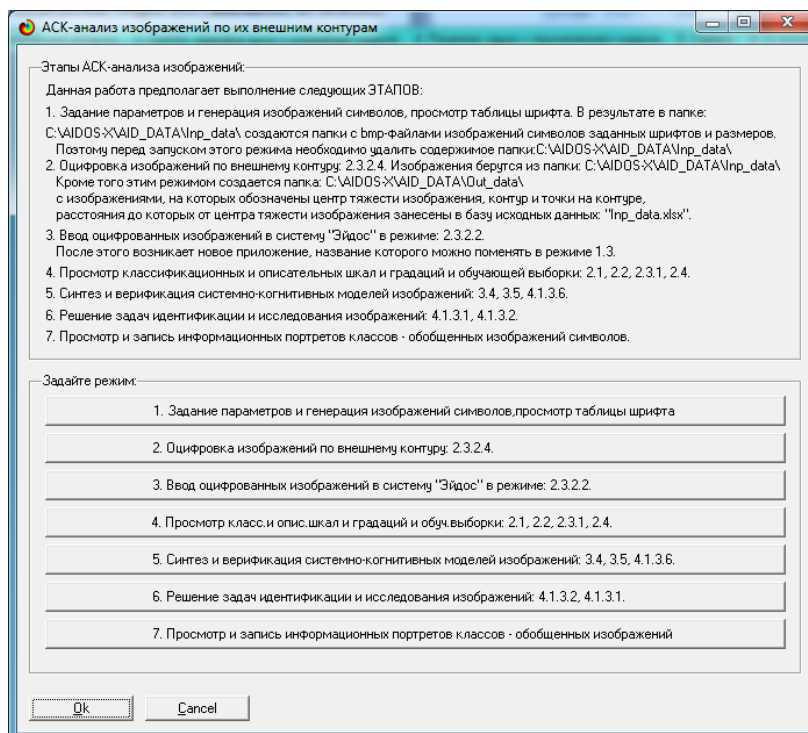


Рисунок 22. Экранная форма режима 4.7. АСК-анализ изображений
Вместо описания возможностей данного режима приведем
его Help (рисунок 23):

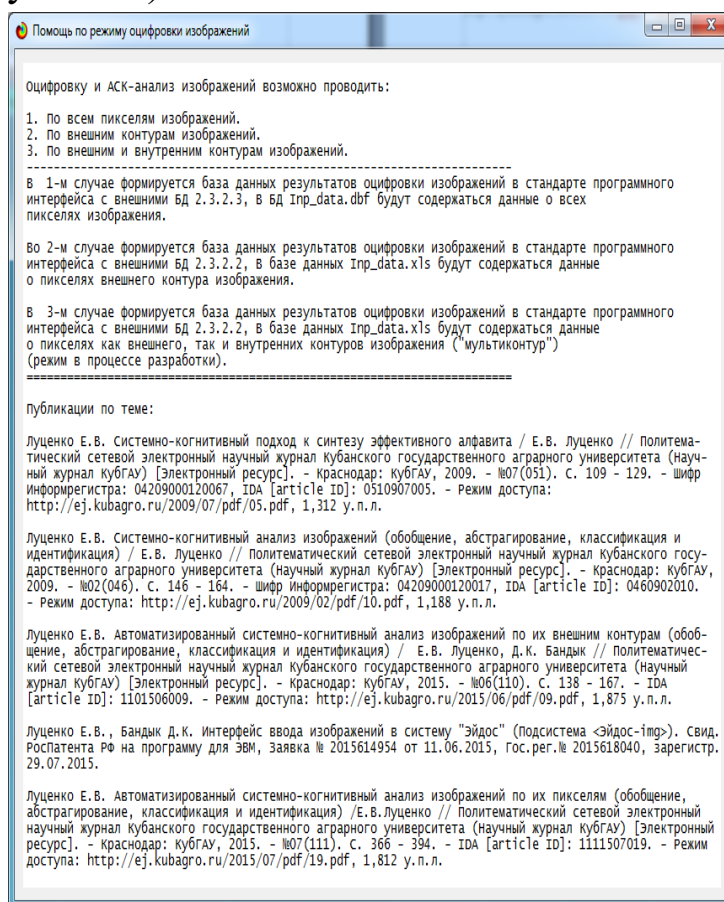


Рисунок 23. Help режима 4.7. АСК-анализ изображений

Экранная форма режима задания параметров, формируемых и отображаемых информационных портретов, приведена на рисунке 24:

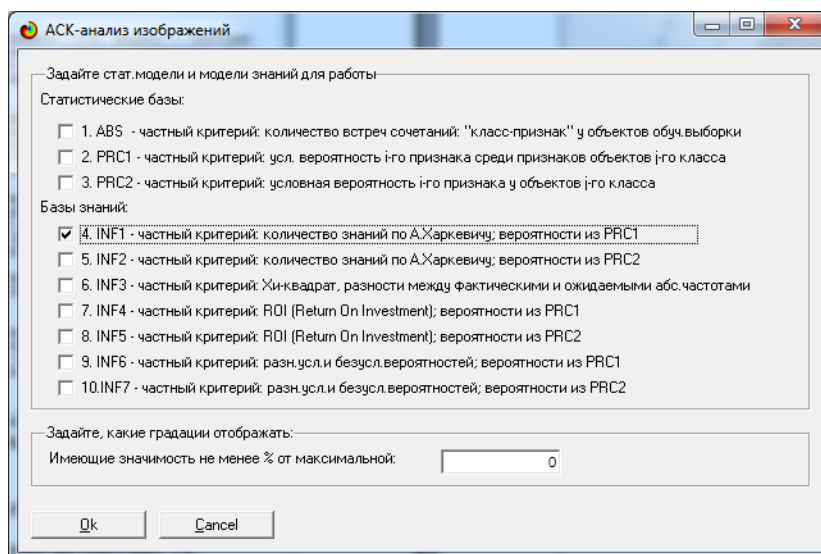


Рисунок 24. Экранная форма режима задания параметров формируемых и отображаемых информационных портретов обобщенных изображений

На рисунке 25 в наглядной графической форме приведена система описательных шкал и градаций, используемая для оцифровки оригинальных изображений листьев, а на рисунке 26 – изображений листьев, усредненных по сортам. На рисунке 25 цветом обозначена ценность градаций описательных шкал для решения задачи идентификации сорта по элементу контура листа. На верхнем рисунке выделена область наиболее значимых элементов контура, на нижнем указаны они все. Для каждого угла поворота радиус-вектора диапазон от минимального до максимальных его значений для конкретных листьев обучающей выборки разделен на 30 интервальных значений – градаций (число интервальных значений задано в диалоге в режиме 2.3.2.2). Решение о количестве интервальных значений задается с учетом объема выборки таким образом, чтобы все они были достаточно широко представлены в ней (по крайней мере по 5 попаданий в каждое интервальное значение каждой шкалы) и не было вообще не представленными. Это значит, что при небольшом

объеме выборки корректно выбирать небольшое количество интервальных значений и соответственно точность модели будет не очень высока, а при увеличении объема выборки есть возможность повысить точность модели. Это одно из следствий известной теоремы Котельникова об отсчетах.

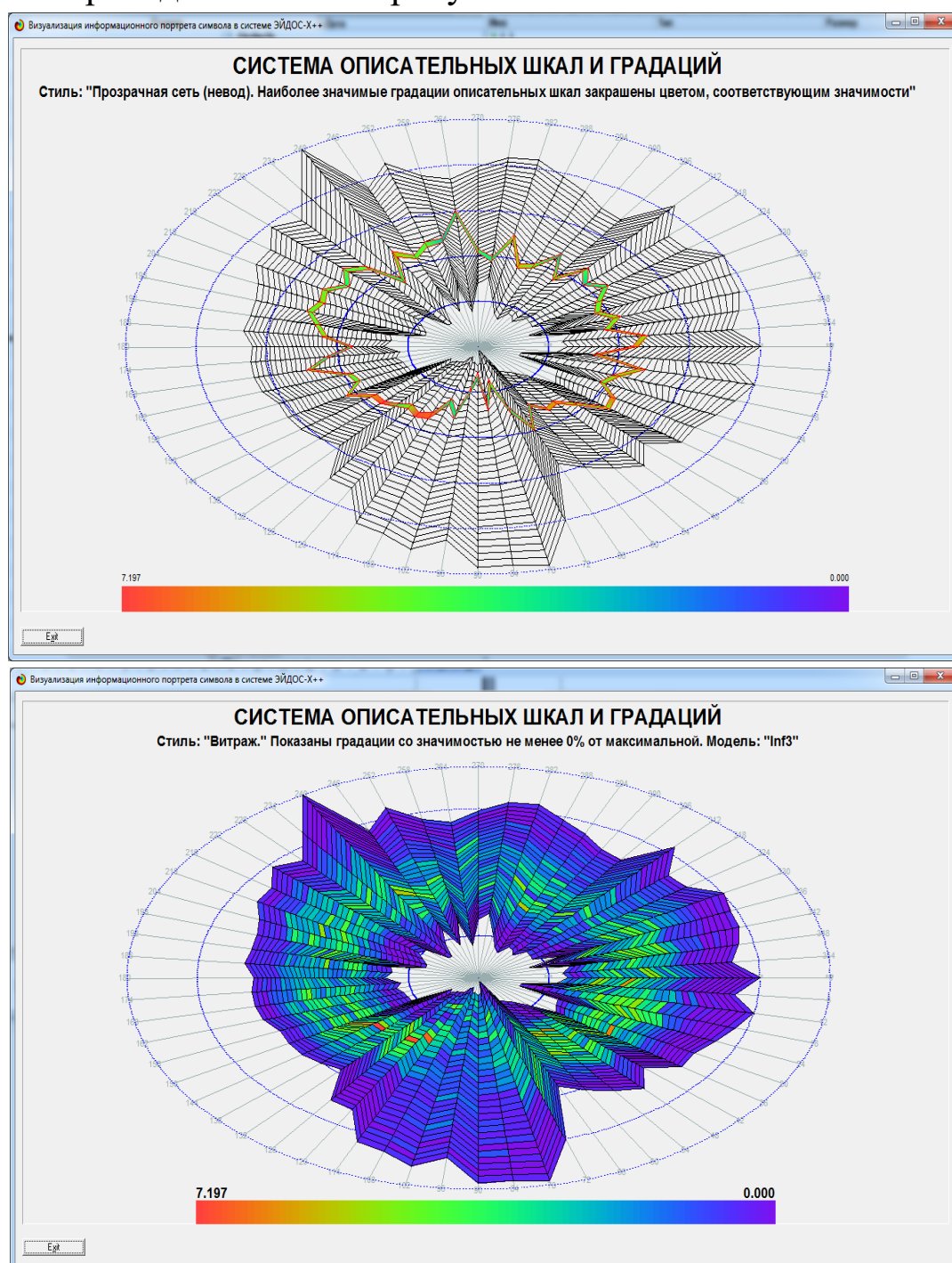
Из сравнения шкал и градаций на рисунках 25 и 26 хорошо видно, на сколько в действительности различаются изображения листьев сортов и на сколько велик уровень шума, вносящий вариабельность в формы конкретных листьев.

Из рисунка 26 видно, что *усредненная форма листьев разных сортов довольно мало отличается друг от друга. Ясно, что чем выше отличие (вариабельность) элементов усредненных контуров по сортам, тем выше ценность этих элементов для того, чтобы отличить эти сорта друг от друга.*

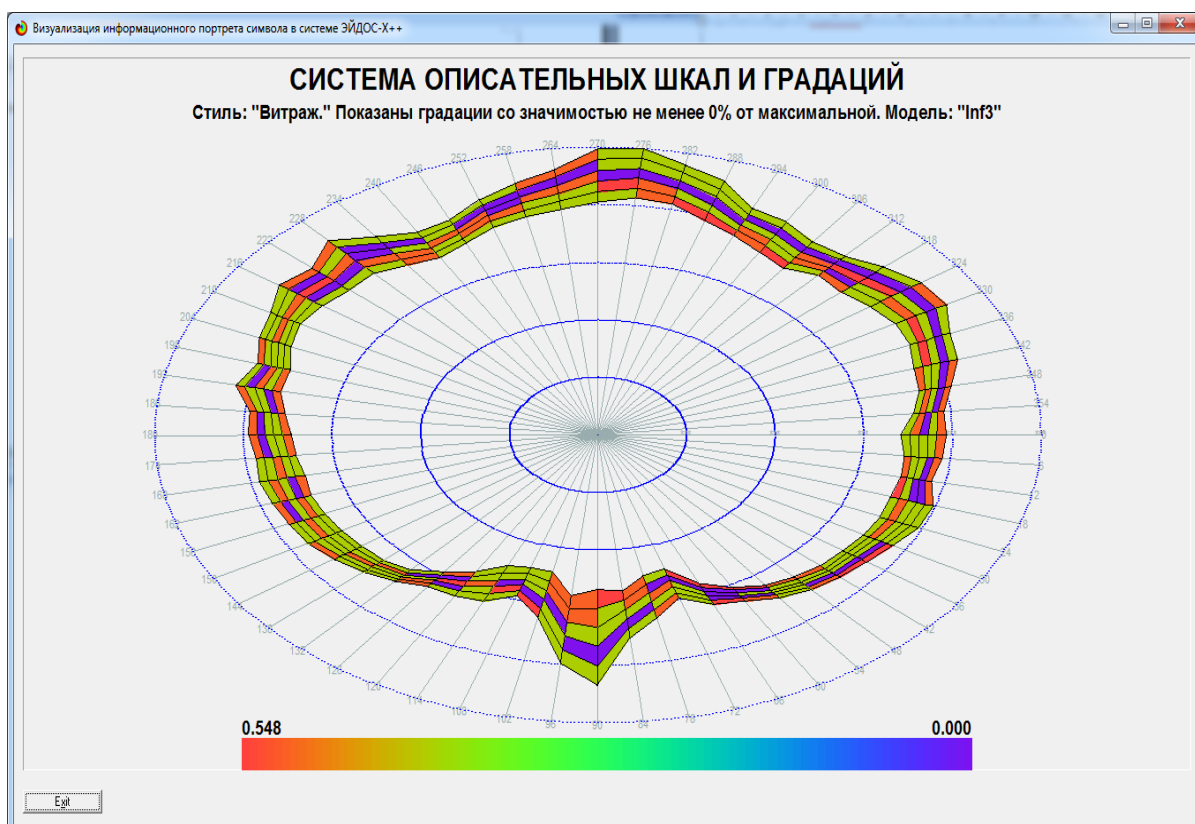
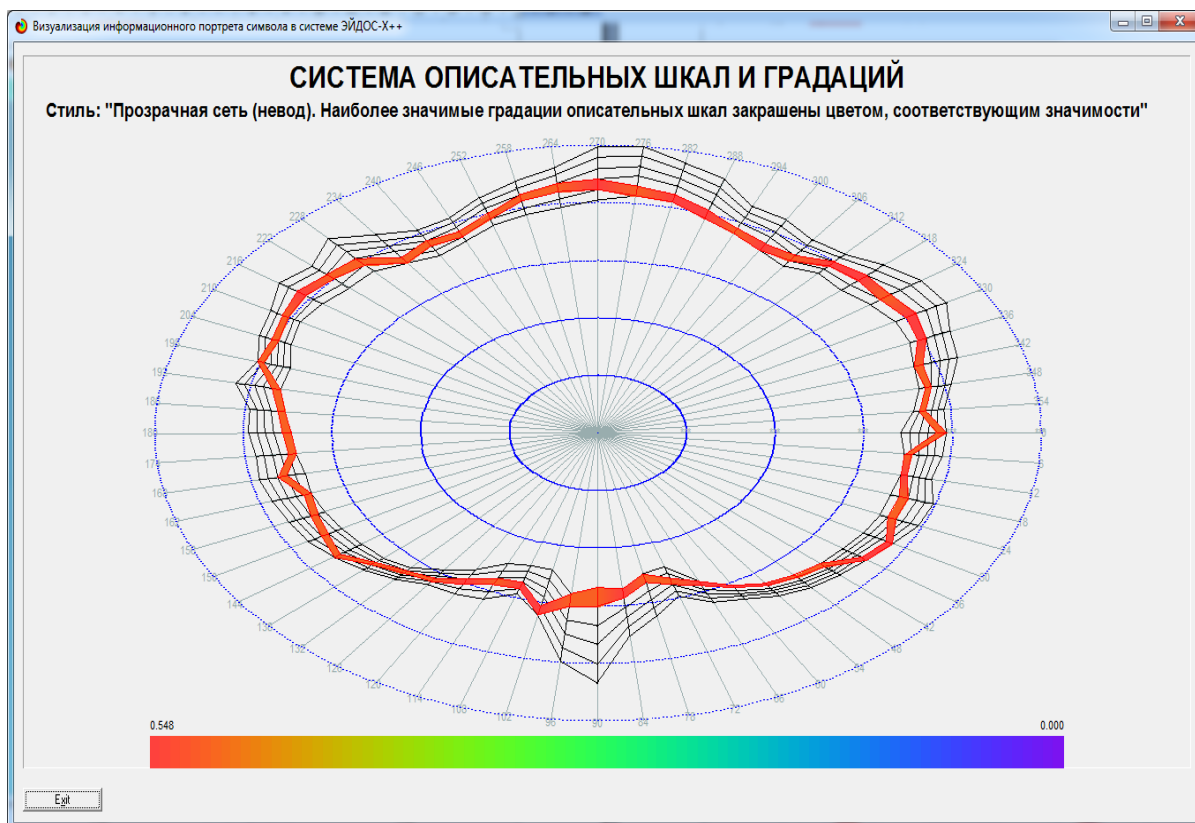
На рисунках 27 приведены информационные портреты обобщенных образов листьев различных сортов, отраженных в модели, построенной по *оригинальным изображениям листьев*, с указанием количества информации в элементе контура о том, что лист с данным элементом принадлежит данному сорту. Количество информации в элементе контура кодируется цветом спектра: максимальное количество информации отображается красным цветом, а минимальное – фиолетовым.

На рисунках 28 приведены информационные портреты обобщенных образов листьев различных сортов, отраженных в модели, построенной по *усредненным по сортам изображениям листьев*, с указанием количества информации в элементе контура о том, что лист с данным элементом принадлежит данному сорту. Количество информации в элементе контура кодируется цветом спектра: максимальное количество информации отображается красным цветом, а минимальное – фиолетовым. Если считать, что форма контура реального листа является суммой идеальной формы листа того сорта, к которому относится лист, и белого

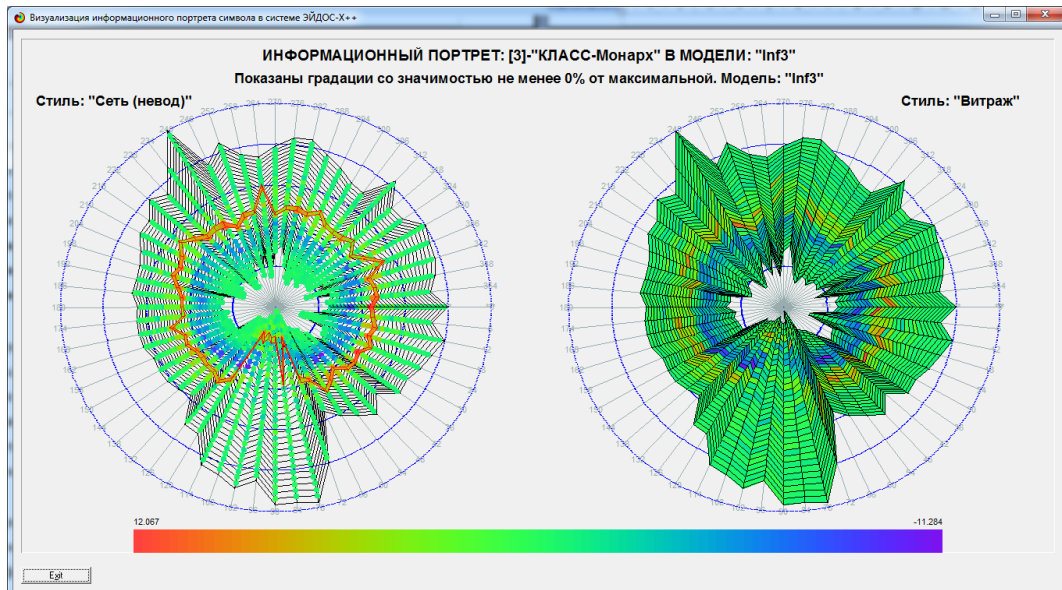
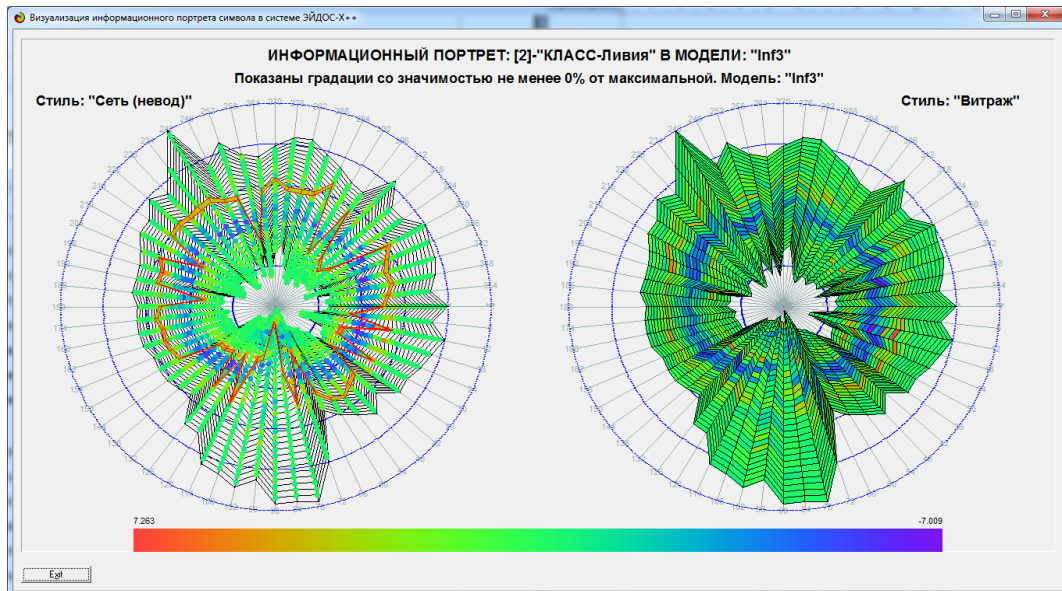
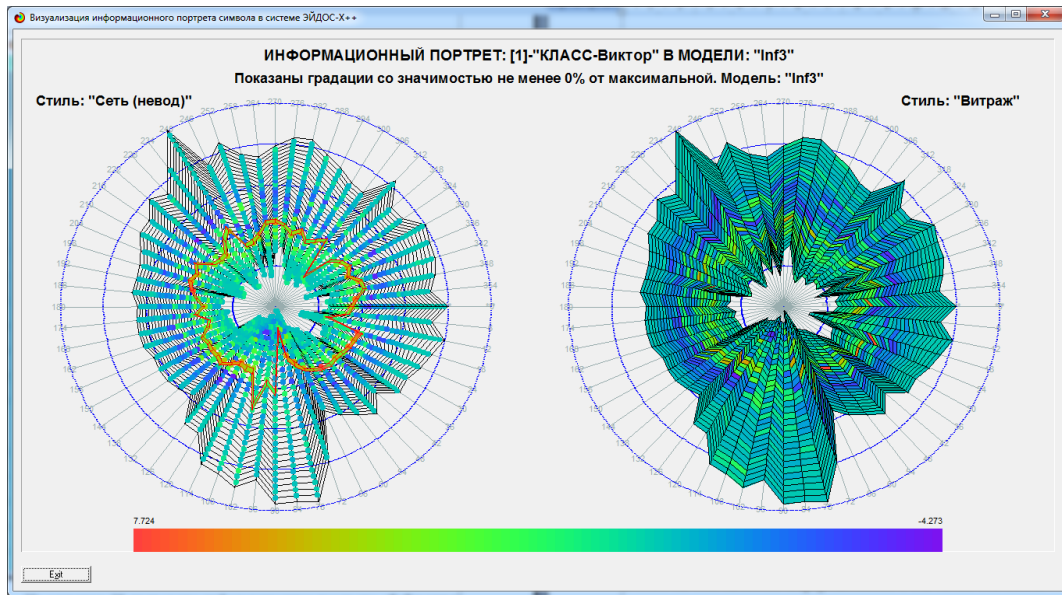
шума, обусловленного случайным воздействием окружающей среды и вносящего искажение в эту форму, то можно предположить, что при усреднении формы листьев по сортам этот шум может быть подавлен. Поэтому исследование информационных портретов усредненной формы листьев разных сортов приведены ниже на рисунках 28.

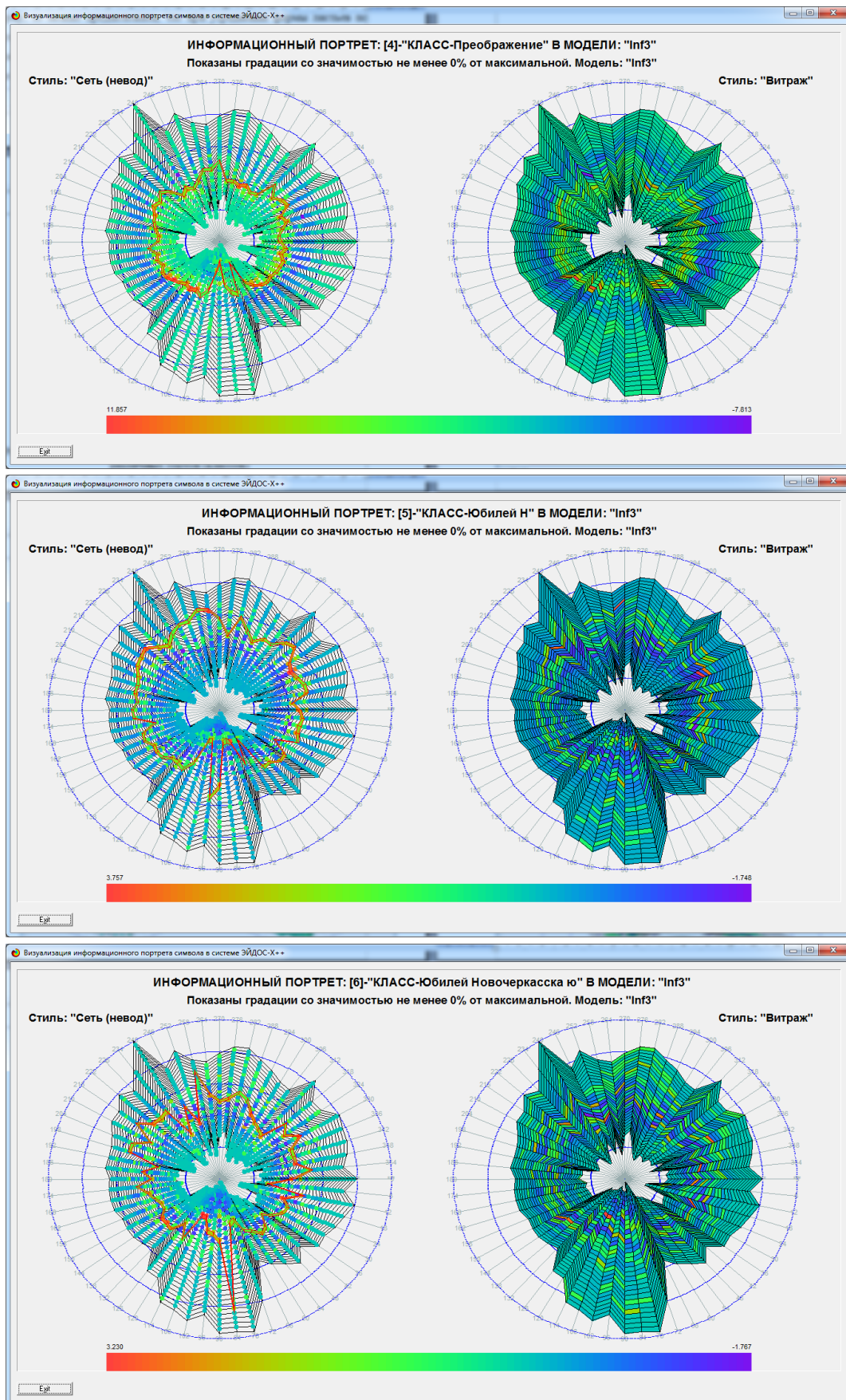


Рисунки 25. Система описательных шкал и градаций, используемая для оцифровки оригинальных изображений листьев

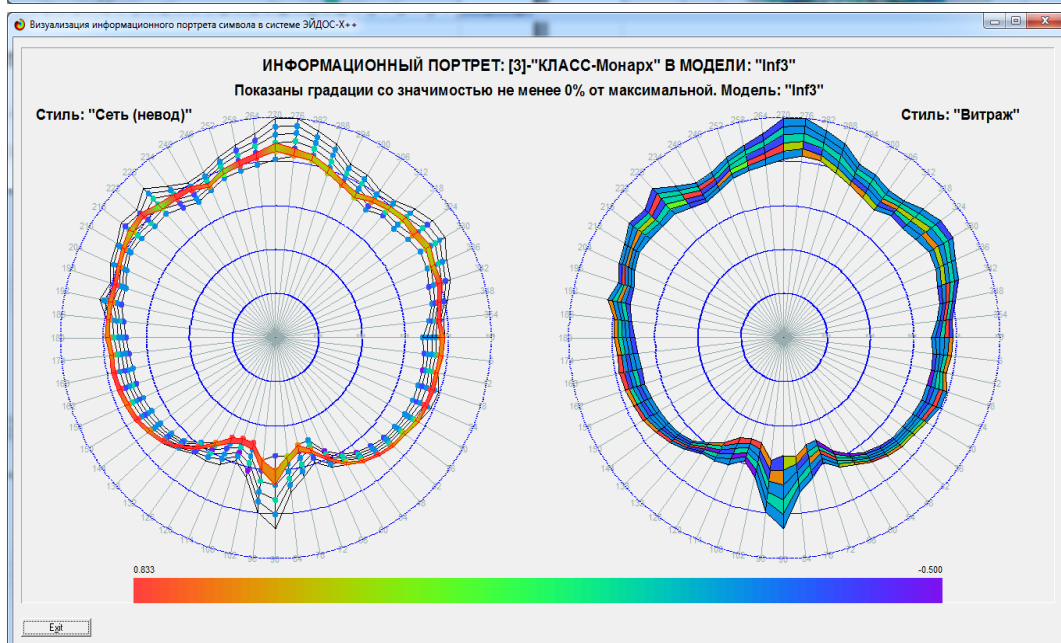
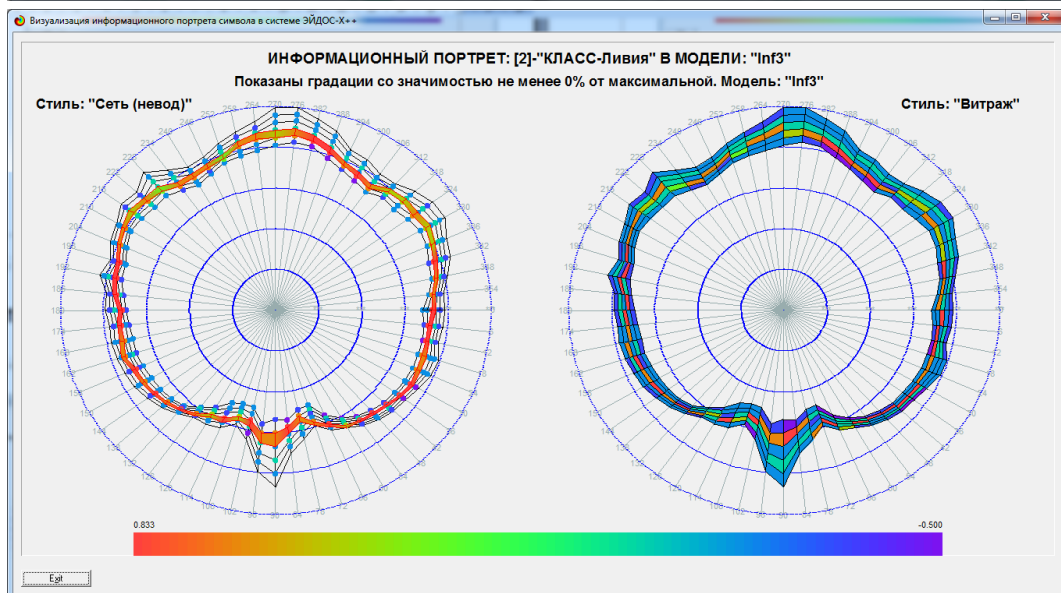
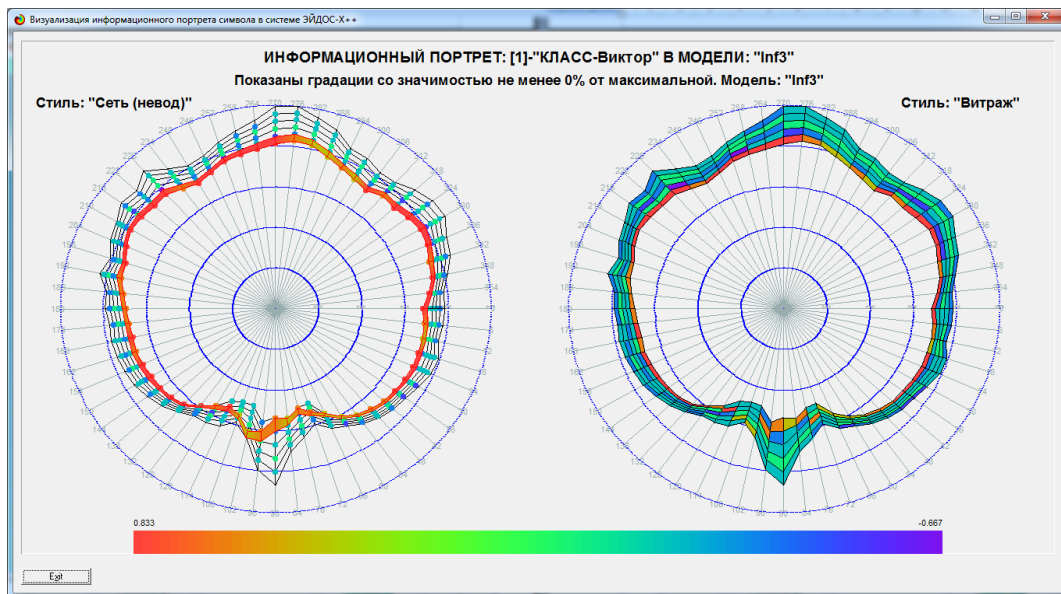


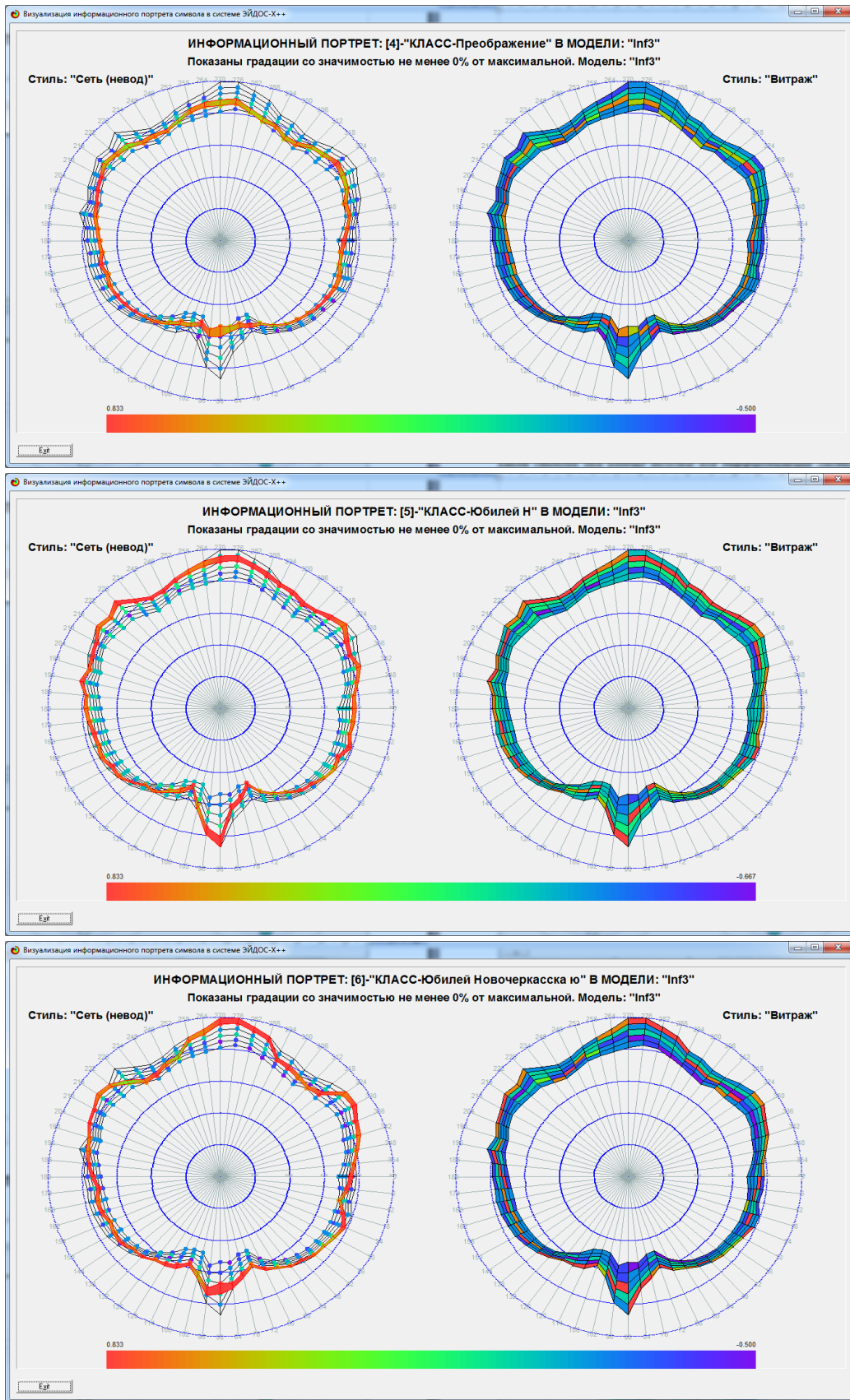
Рисунки 26. Система описательных шкал и градаций, используемая для оцифровки усредненных по сортам изображений листьев





Рисунки 27. Информационные портреты обобщенных образов листьев различных сортов (классов)





Рисунки 28. Информационные портреты обобщенных образов листьев различных сортов, построенные по модели INF1 для усредненных по сортам изображений листьев

Из портретов на рисунках 27 и 28 видно, какое количество информации о принадлежности листа к каждому из сортов содержат различные элементы контура, а из рисунков 25 и 26 видно, в какой степени эти элементы контура вносят различие между сортами, т.е. в какой степени они вообще полезны для дифференциации листьев по сортам. По сути это означает, что на рисунках 27 и 28 приведены редуцированные и нередуцированные когнитивные функции, причем редуцированные когнитивные функции по сути представляют нам вид идеального, т.е. не зашумленного случайным воздействием окружающей среды, контура листа каждого сорта.

Различные варианты частных и интегральных критериев, применяемых в системе «Эйдос», описаны в работе [7] и других работах по АСК-анализу [12, 13].

По-видимому, при решении задачи идентификации сорта по образцу листа есть смысл обращать основное внимание именно на эти наиболее информативные и значимые элементы контуров.

В системе «Эйдос» реализован фильтр, в результате действия которого на изображениях будут показаны лишь те элементы контуров изображений, которые имеют значимость не менее заданной пользователем величины.

Примеры изображений шкал и информационных портретов с фильтром, ограничивающим отображение градаций шкал с низким уровнем значимости, приведены на рисунках 29 и 30:

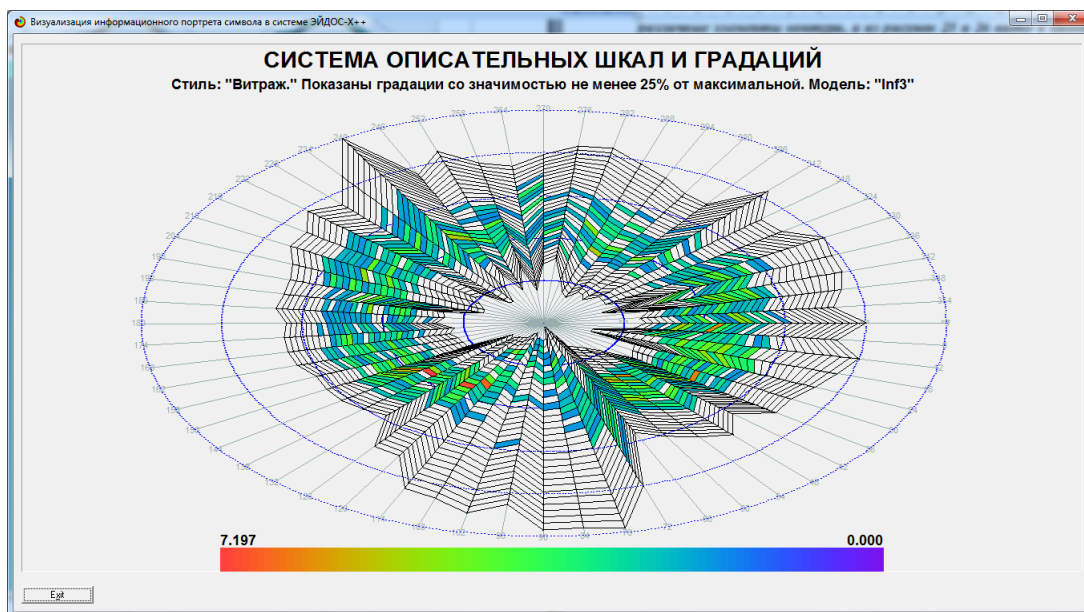


Рисунок 29. Система описательных шкал и градаций, используемая для оцифровки оригинальных изображений листьев, с фильтром, ограничивающим отображение градаций шкал с низким уровнем значимости

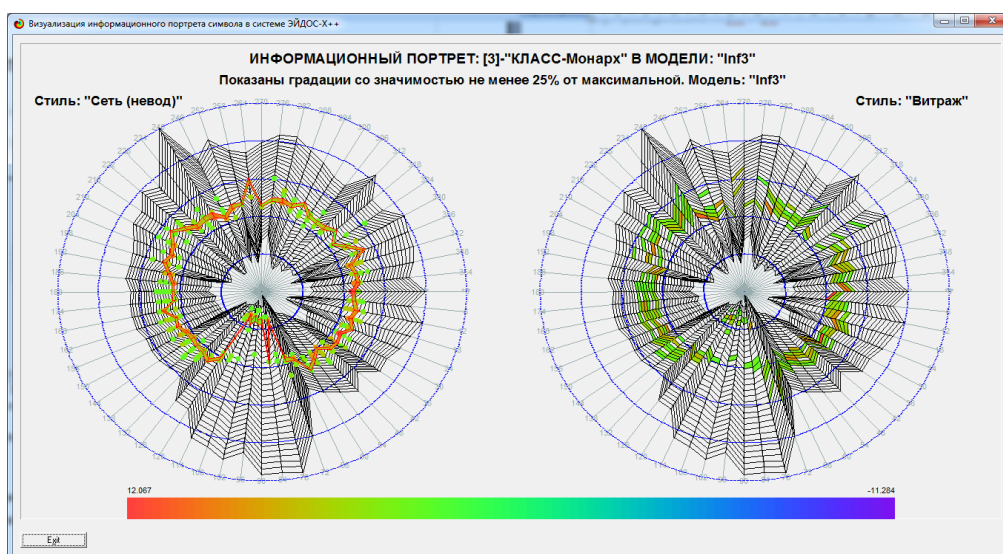


Рисунок 30. Информационный портрет обобщенного образа листьев сорта «Монарх», построенная в модели INF3 на основе оригинальных изображений листьев, с фильтром, ограничивающим отображение градаций шкал, нехарактерных (малоинформативных) для данного сорта

Задача 4: сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу.

Результаты идентификации листьев с сортами и сортов с листьями приводятся в многочисленных выходных формах (рисунок 31), из которых мы в связи с ограниченностью объема работы приведем лишь первые 2 (рисунки 32 и 33):

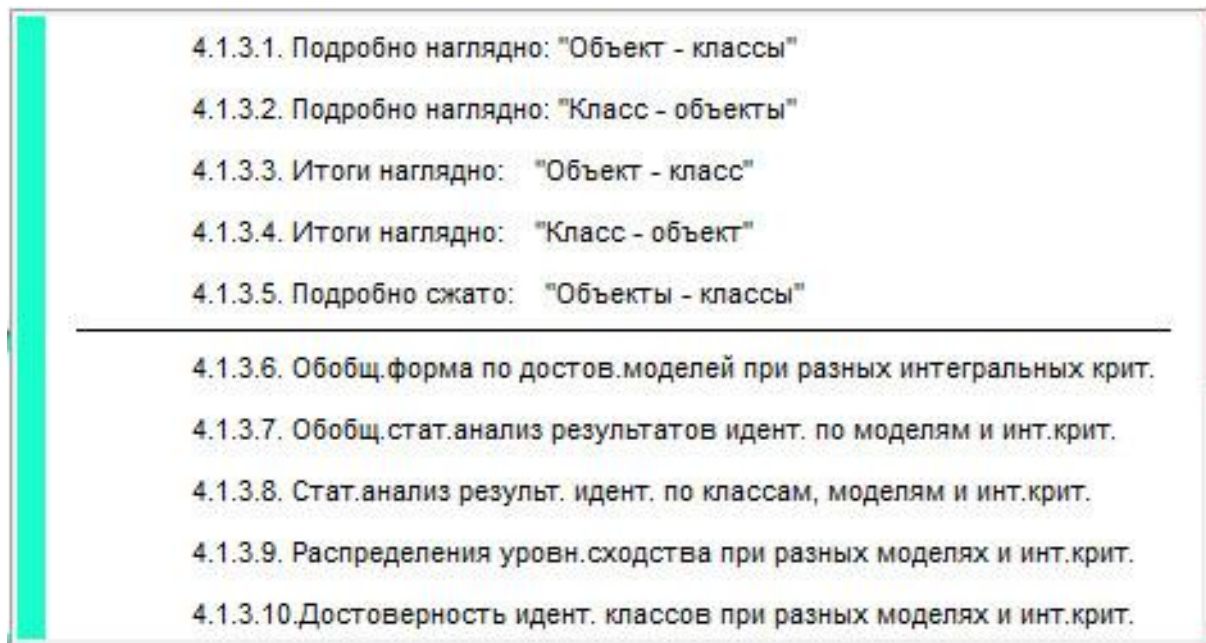


Рисунок 31. Выходные формы с результатами идентификации

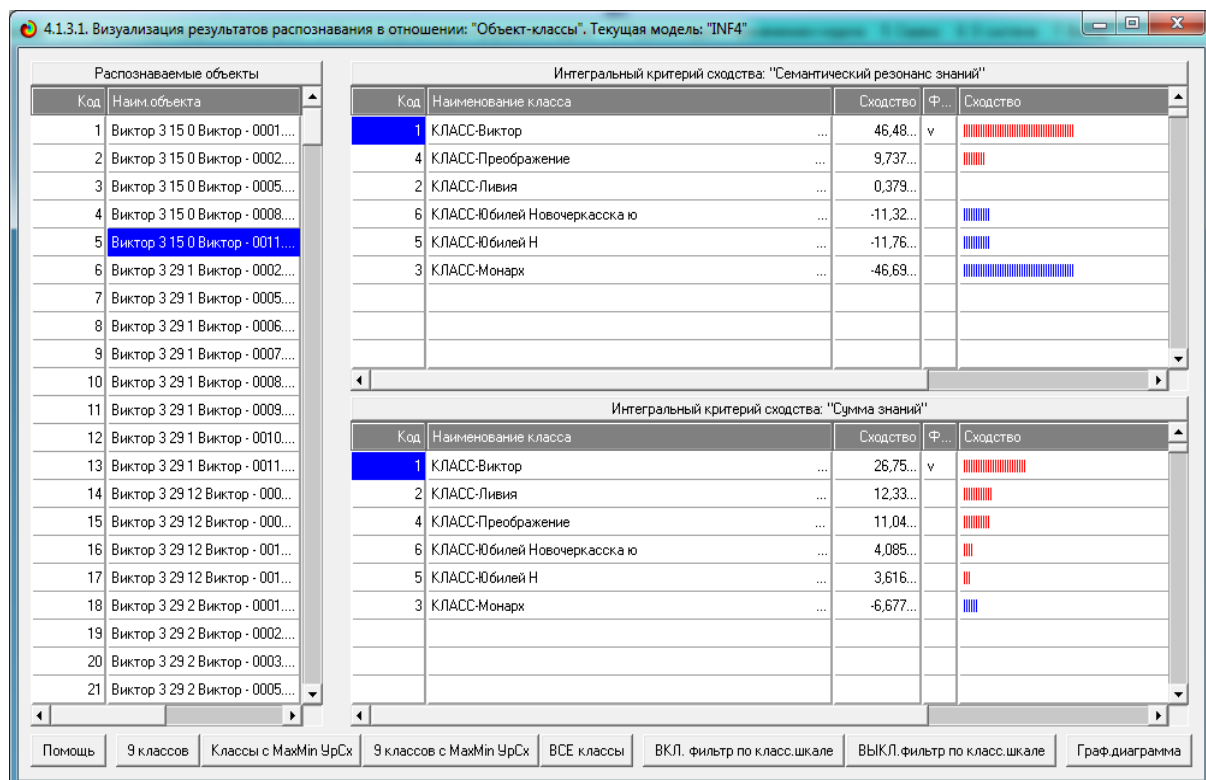


Рисунок 32. Результаты идентификации листа с сортами

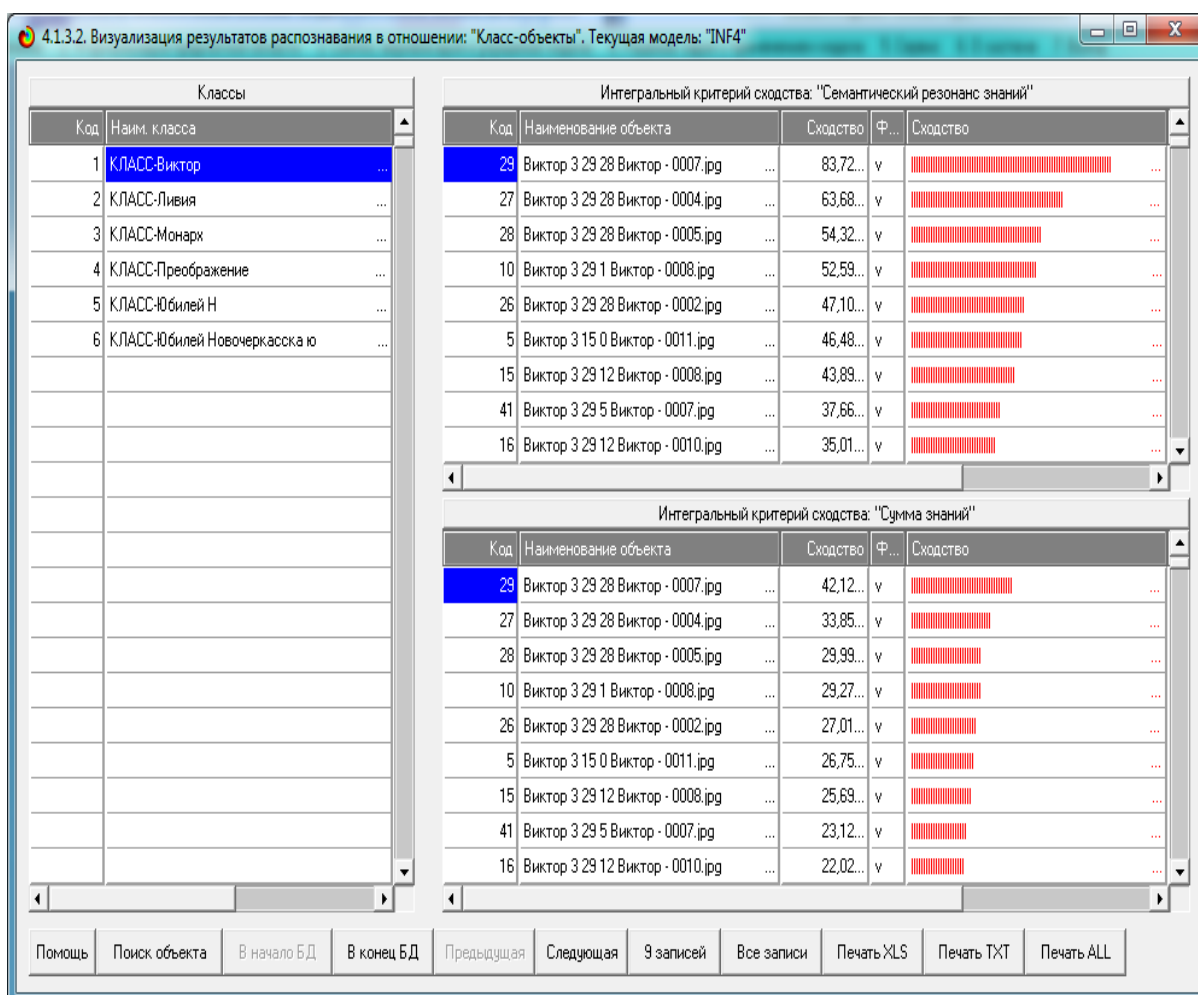


Рисунок 33. Результаты идентификации сорта с листьями

Объект, относящийся к классу, практически всегда имеет максимальный уровень сходства с ним, по сравнению с другими объектами. Мы видим, что *результаты весьма разумные и по своей достоверности соответствуют уровню эксперта в данной предметной области.*

Задача 5: количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов (решение задачи классификации).

Для решения этой задачи необходимо запустить режим 4.2.2.1, а затем 4.2.2.2 (рисунки 31 и 32):

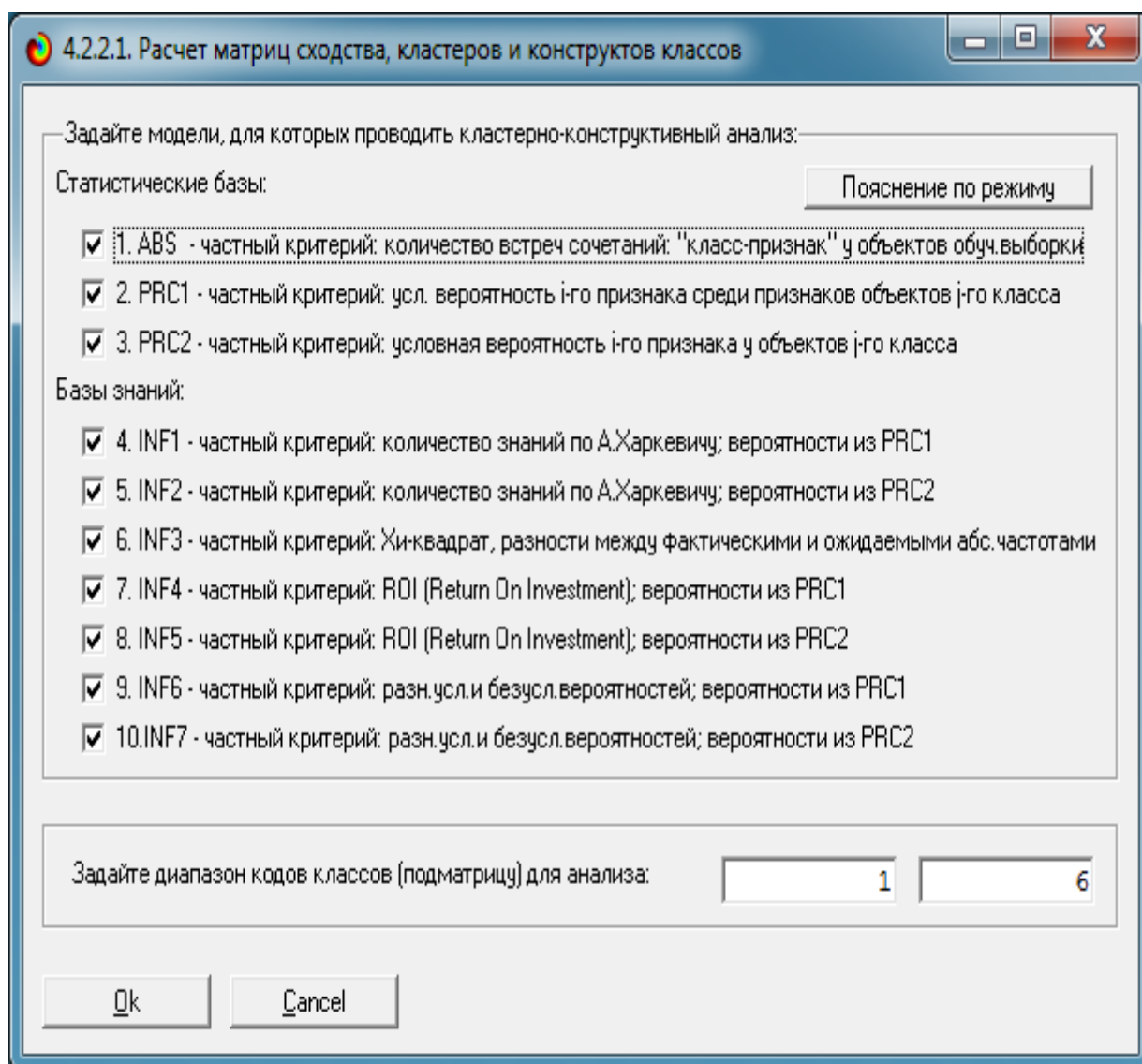


Рисунок 34. Экранная форма запуска режима кластерно-конструктивного анализа

При кластерном анализе наиболее сходные классы объединяются в группы, которые называются кластерами. При конструктивном анализе наиболее сильно отличающиеся кластеры рассматриваются как смысловые полюса конструктора, а другие классы и кластеры с промежуточными смысловыми градациями образуют смысловой спектр между полюсами. Кластерно-конструктивный анализ предложен проф. Е.В. Луценко в 2002 году как одна из возможностей АСК-анализа и реализован в системе «Эйдос» [13].

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 1 "КЛАСС-Виктор" в модели: 4 "INF1"

| Код | Наименование класса | № | Код класса | Наименование класса | Сходство |
|-----|------------------------------|---|------------|------------------------------|----------|
| 1 | КЛАСС-Виктор | 1 | 1 | КЛАСС-Виктор | 100.000 |
| 2 | КЛАСС-Ливия | 2 | 4 | КЛАСС-Преобразование | 12.605 |
| 3 | КЛАСС-Монарх | 3 | 2 | КЛАСС-Ливия | 9.533 |
| 4 | КЛАСС-Преобразование | 4 | 6 | КЛАСС-Юбилей Новочеркасска ю | -8.483 |
| 5 | КЛАСС-Юбилей Н | 5 | 5 | КЛАСС-Юбилей Н | -15.456 |
| 6 | КЛАСС-Юбилей Новочеркасска ю | 6 | 3 | КЛАСС-Монарх | -25.753 |

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ.фильтр по кл.шкале ВЫКЛ.фильтр по кл.шкале Вписать в окно Показать ВСЕ

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструкт класса: 1 "КЛАСС-Виктор" в модели: 7 "INF4"

| Код | Наименование класса | № | Код класса | Наименование класса | Сходство |
|-----|------------------------------|---|------------|------------------------------|----------|
| 1 | КЛАСС-Виктор | 1 | 1 | КЛАСС-Виктор | 100.000 |
| 2 | КЛАСС-Ливия | 2 | 4 | КЛАСС-Преобразование | -1.014 |
| 3 | КЛАСС-Монарх | 3 | 6 | КЛАСС-Юбилей Новочеркасска ю | -4.383 |
| 4 | КЛАСС-Преобразование | 4 | 5 | КЛАСС-Юбилей Н | -7.042 |
| 5 | КЛАСС-Юбилей Н | 5 | 2 | КЛАСС-Ливия | -7.326 |
| 6 | КЛАСС-Юбилей Новочеркасска ю | 6 | 3 | КЛАСС-Монарх | -19.671 |

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ.фильтр по кл.шкале ВЫКЛ.фильтр по кл.шкале Вписать в окно Показать ВСЕ

Рисунки 35. Табличные экранные формы с результатами кластерно-конструктивного анализа в когнитивных моделях INF1 и INF4

Из сравнения первой и второй экранных форм на рисунке 31 видно, что в более достоверной модели INF4 сорта сильнее отличаются друг от друга, что соответствует интуитивным ожиданиям экспертов, основанным на их опыте и профессиональной компетенции. На рисунке 33 результаты сравнения друг с другом обобщенных образов листьев различных

сортов винограда приведены в графической форме когнитивной диаграммы:

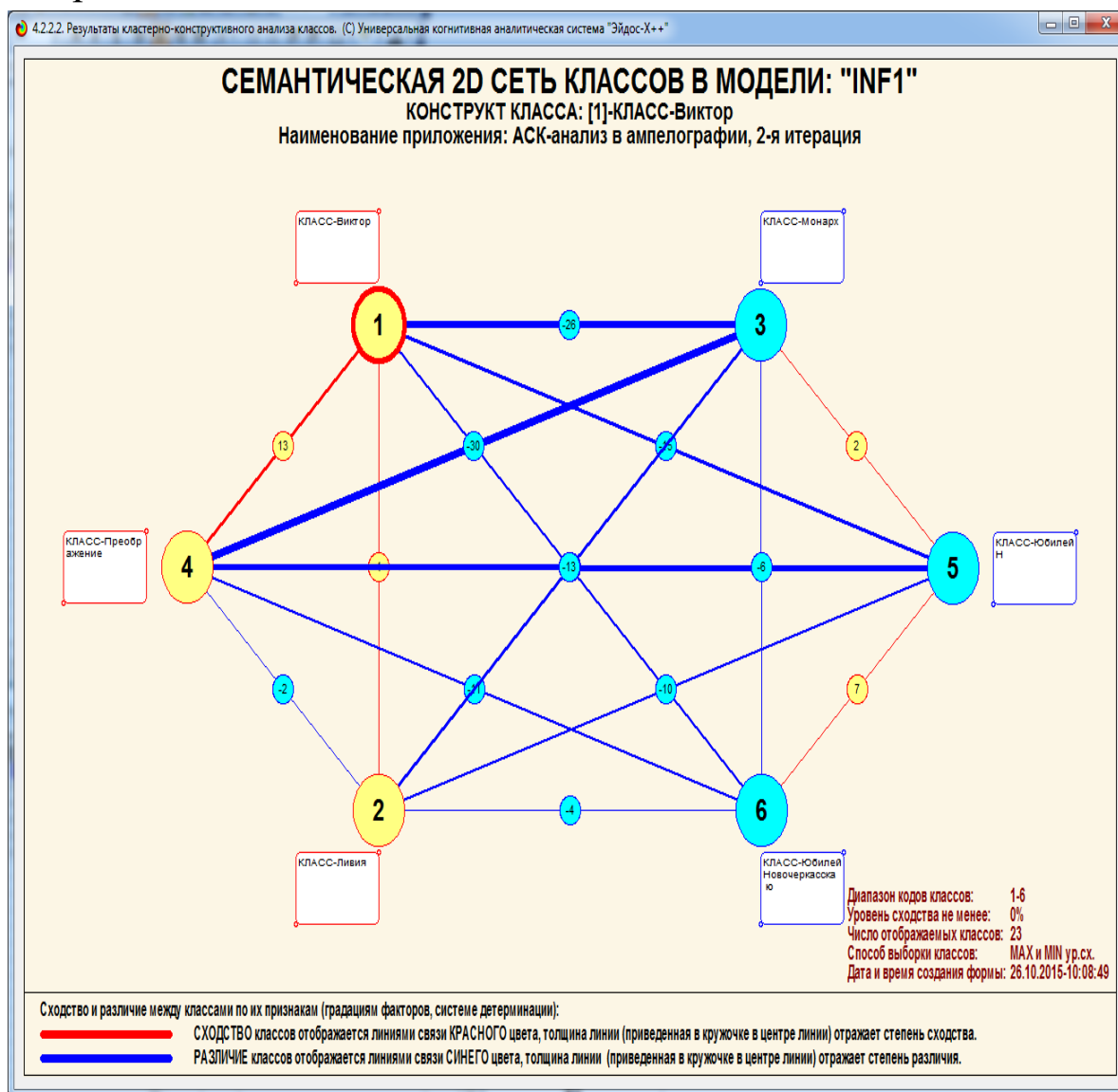


Рисунок 36. Результаты сравнения друг с другом обобщенных образов классов (идеальных листьев различных сортов)

Сорта «Виктор» и «Преобразование» являются сибсами – родственниками, поэтому у них сходная форма листьев, что количественно строго объективно и определено системой непосредственно на основе эмпирических данных без участия экспертов. Сорт «Юбилей Новочеркаска» также является родственником этих двух сортов, однако его форма листьев отличается от них. Сорта Ливия К и Монарх (ныне назван

Хризолитом) получены путем скрещивания различных исходных форм винограда [14-17], поэтому естественно, что они отличаются друг от друга по форме листьев.

SWOT-анализ изображений. Характерные и нехарактерные признаки контуров обобщенных изображений листьев по сортам можно получить в форме таблиц и графических диаграмм в режиме 4.4.8 (рисунки 34):

4.4.8. Количественный автоматизированный SWOT-анализ классов средствами АСК-анализа в системе "Эйдос"

Выбор класса, соответствующего будущему состоянию объекта управления

| Код | Наименование класса |
|-----|-----------------------------|
| 2 | КЛАСС-Ливия |
| 3 | КЛАСС-Монарх |
| 4 | КЛАСС-Преображение |
| 5 | КЛАСС-Юбилей Н |
| 6 | КЛАСС-Юбилей Новочеркаска ю |

SWOT-анализ класса: 6 "КЛАСС-Юбилей Новочеркаска ю" в модели: 7 "INF4"

Способствующие факторы и сила их влияния

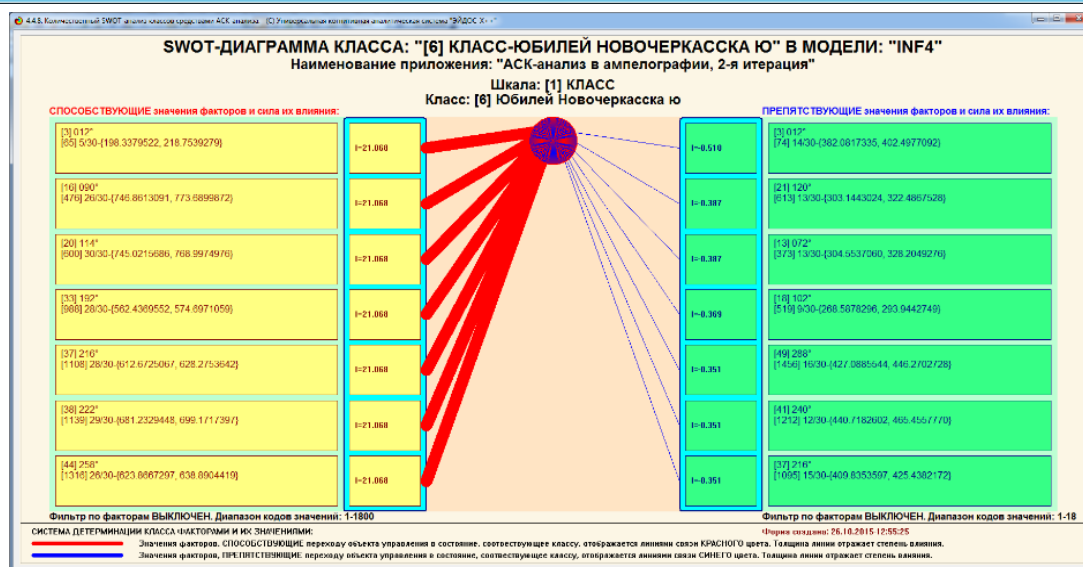
| Код | Наименование фактора и его интервального значения | Сила влияния |
|-------|---|--------------|
| 65 | 012°-5/30-(198.3379522, 218.7539279) | 21.068 |
| 476 | 090°-26/30-(746.8613091, 773.6899872) | 21.068 |
| 600 | 114°-30/30-(745.0215686, 768.9974976) | 21.068 |
| 988 | 192°-28/30-(562.4369552, 574.6971059) | 21.068 |
| 11... | 216°-28/30-(612.6725067, 628.2753642) | 21.068 |
| 11... | 222°-29/30-(681.2329448, 699.1717397) | 21.068 |
| 13... | 258°-26/30-(623.8667297, 638.8904419) | 21.068 |
| 13... | 270°-27/30-(657.0897909, 672.6378235) | 21.068 |
| 14... | 276°-30/30-(743.9562917, 759.9575806) | 21.068 |
| 14... | 282°-28/30-(706.3435837, 725.3665477) | 21.068 |
| 16... | 324°-26/30-(583.6925151, 600.3370199) | 21.068 |
| 142 | 024°-22/30-(538.2998322, 553.1301921) | 10.034 |

Препятствующие факторы и сила их влияния

| Код | Наименование фактора и его интервального значения | Сила влияния |
|-------|---|--------------|
| 74 | 012°-14/30-(382.0817335, 402.4977092) | -0.510 |
| 613 | 120°-13/30-(303.1443024, 322.4867528) | -0.387 |
| 373 | 072°-13/30-(304.5537060, 328.2049276) | -0.387 |
| 519 | 102°-9/30-(268.5878296, 293.9442749) | -0.369 |
| 14... | 288°-16/30-(427.0885544, 446.2702728) | -0.351 |
| 12... | 240°-12/30-(440.7182602, 465.4557770) | -0.351 |
| 10... | 216°-15/30-(409.8353597, 425.4382172) | -0.351 |
| 736 | 144°-16/30-(323.1285401, 339.0849325) | -0.351 |
| 581 | 114°-11/30-(289.4789175, 313.4548465) | -0.351 |
| 552 | 108°-12/30-(315.2997190, 340.0186020) | -0.351 |
| 16... | 324°-18/30-(450.5364767, 467.1809815) | -0.331 |
| 15... | 312°-13/30-(385.2114288, 406.0317174) | -0.331 |

ВКЛЮЧИТЬ фильтр по фактору ВЫКЛЮЧИТЬ фильтр по фактору

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 Нейрон SWOT-диаграмма Интегральная когнитивная карта



Рисунки 37. Характерные и нехарактерные признаки контуров обобщенного изображения листьев по сорту «Юбилей Новочеркаска»

Парето-оптимизация (абстрагирование). Не все описательные шкалы и градации имеют одинаковую ценность для идентификации изображений. Наиболее ценные могут использоваться для решения задач, а наименее ценные вообще могут не учитываться в моделях практически без ущерба для их достоверности (рисунок 35):

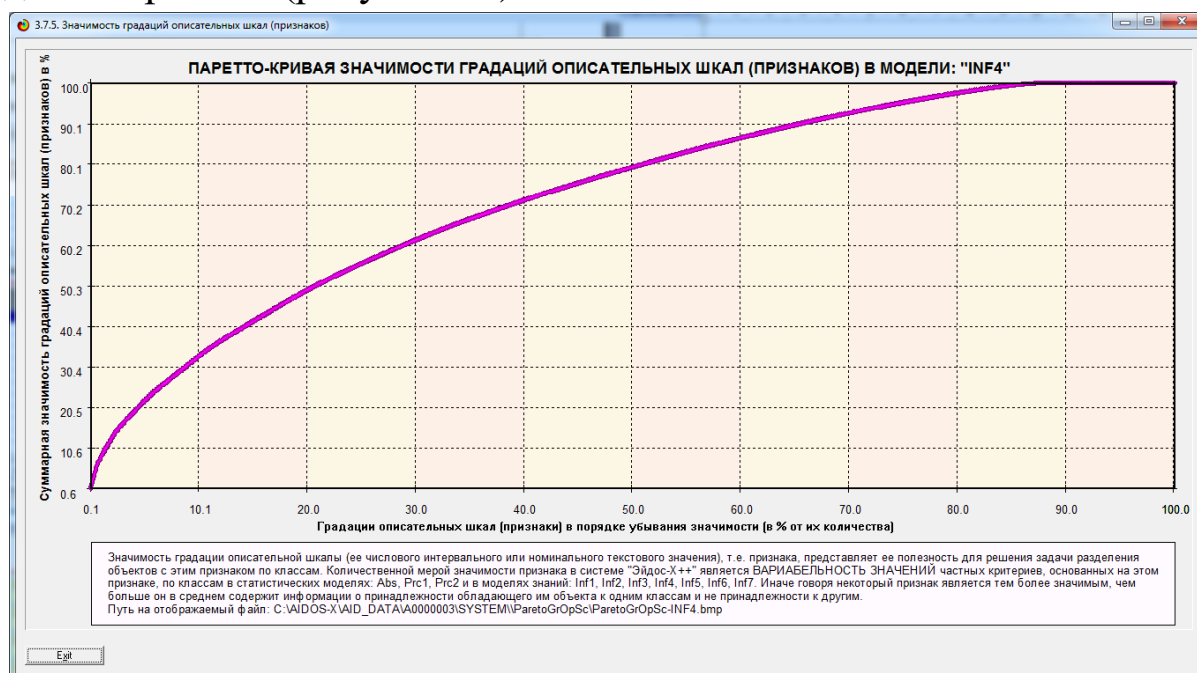


Рисунок 38. Общий вид Парето-кривой ценности градаций описательных шкал (элементов контуров изображений) для идентификации конкретных изображений с обобщенными образами классов

Расшифровка наименований градаций описательных шкал, ранжированных в порядке убывания их ценности, дается в дополнительных таблицах, которые в данной статье приводить нецелесообразно.

Выводы

В главе рассмотрено применение автоматизированного системно-когнитивного анализа (АСК-анализ), его

математической модели – системной теории информации и программного инструментария – интеллектуальной системы «Эйдос» для решения ряда задач ампелографии: 1) оцифровка сканированных изображений листьев и создание их математических моделей; 2) формирование математических моделей конкретных листьев с применением теории информации; 3) формирование моделей обобщенных образов листьев различных сортов; 4) сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу; 5) количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов. Предложен новый подход к оцифровке изображений листьев, основанный на использовании полярной системы координат, центра тяжести изображения и его внешнего контура. Перед оцифровкой изображений могут применяться их преобразования, стандартизирующие положение изображений, их размеры и угол поворота. Поэтому результаты оцифровки и АСК-анализа изображений могут быть инвариантны (независимы) относительно их положения, размеров и поворота. Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о сорте, включающее как информацию об истинной форме листа данного сорта (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным воздействием окружающей среды. Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение сигнала об истинной форме листа каждого сорта на основе ряда зашумленных конкретных примеров листьев данного сорта. Таким образом, создается один образ формы листа каждого сорта, независящий от их конкретных реализаций, т.е. «Эйдос»

этих изображений (в смысле Платона) - прототип или архетип (в смысле Юнга) изображений.

Материалы данной главы могут быть использованы при преподавании дисциплин: интеллектуальные системы; инженерия знаний и интеллектуальные системы; интеллектуальные технологии и представление знаний; представление знаний в интеллектуальных системах; основы интеллектуальных систем; введение в нейроматематику и методы нейронных сетей; основы искусственного интеллекта; интеллектуальные технологии в науке и образовании; управление знаниями; автоматизированный системно-когнитивный анализ и интеллектуальная система «Эйдос»; которые первый автор ведет в настоящее время⁷, а также и в других дисциплинах, связанных с преобразованием данных в информацию, а ее – в знания и применением этих знаний для решения задач идентификации, прогнозирования, принятия решений и исследования моделируемой предметной области (а это практически все дисциплины во всех областях науки). В частности, с применением предложенной технологии АСК-анализа изображений по их контурам могут быть модифицированы лабораторные работы 1 и 3 [19].

Этим и другим применениям должно способствовать и то, что система «Эйдос» находится в полном открытом бесплатном доступе (с открытыми исходными текстами) на сайте первого автора по адресу: <http://lc.kubagro.ru/aidos/Aidos-X.htm>.

Запланировано совершенствование методологии, технологии, методики и программного инструментария АСК-анализа изображений путем развития, применения и обобщения описанного в данной работе подхода, что будет способствовать развитию науки ампелографии и ее ветвей – ампелометрии и ампелологии.

⁷ http://lc.kubagro.ru/My_training_schedule.doc

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Луценко Е.В. Системно-когнитивный анализ изображений (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №02(046). С. 146 – 164. – Шифр Информрегистра: 0420900012\0017, IDA [article ID]: 0460902010. – Режим доступа: <http://ej.kubagro.ru/2009/02/pdf/10.pdf>, 1,188 у.п.л.
2. Луценко Е.В. Системно-когнитивный подход к синтезу эффективного алфавита / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №07(051). С. 109 – 129. – Шифр Информрегистра: 0420900012\0067, IDA [article ID]: 0510907005. – Режим доступа: <http://ej.kubagro.ru/2009/07/pdf/05.pdf>, 1,312 у.п.л.
3. Луценко Е.В. Автоматизированный системно-когнитивный анализ изображений по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №06(110). С. 138 – 167. – IDA [article ID]: 1101506009. – Режим доступа: <http://ej.kubagro.ru/2015/06/pdf/09.pdf>, 1,875 у.п.л.
4. Луценко Е.В., Бандык Д.К. Интерфейс ввода изображений в систему "Эйдос" (Подсистема «Эйдос-img»). Свид. РосПатента РФ на программу для ЭВМ, Заявка № 2015614954 от 11.06.2015, Гос.рег.№ 2015618040, зарегистр. 29.07.2015. – Режим доступа: <http://lc.kubagro.ru/aidos/2015618040.jpg>.
5. Луценко Е.В. Автоматизированный системно-когнитивный анализ изображений по их пикселям (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №07(111). С. 334 – 362. – IDA [article ID]: 1111507019. – Режим доступа: <http://ej.kubagro.ru/2015/07/pdf/19.pdf>, 1,812 у.п.л.
6. Луценко Е.В. СК-анализ и система "Эйдос" в свете философии Платона / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №01(045). С. 91 – 100. – Шифр Информрегистра: 0420900012\0010, IDA [article ID]: 0450901008. – Режим доступа: <http://ej.kubagro.ru/2009/01/pdf/08.pdf>, 0,625 у.п.л.
7. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 у.п.л.
8. Луценко Е.В. Повышение адекватности спектрального анализа личности по астросоциотипам путем их разделения на типичную и нетипичную части / Е.В. Луценко, А.П. Трунев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2008. – №02(036). С. 153 – 174.

– Шифр Информрегистра: 0420800012\0017, IDA [article ID]: 0360802010. – Режим доступа: <http://ej.kubagro.ru/2008/02/pdf/10.pdf>, 1,375 у.п.л.

9. Луценко Е.В. Повышение качества моделей «knowledge management» путем разделения классов на типичную и нетипичную части / Е.В. Луценко, Е.А. Лебедев, В.Н. Лаптев // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №10(054). С. 78 – 93. – Шифр Информрегистра: 0420900012\0109, IDA [article ID]: 0540910005. – Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/05.pdf>, 1 у.п.л.

10. Луценко Е.В. 30 лет системе «Эйдос» – одной из старейших отечественных универсальных систем искусственного интеллекта, широко применяемых и развивающихся и в настоящее время / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №10(054). С. 48 – 77. – Шифр Информрегистра: 0420900012\0110, IDA [article ID]: 0540910004. – Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/04.pdf>, 1,875 у.п.л.

11. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос-Х++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №09(083). С. 328 – 356. – IDA [article ID]: 0831209025. – Режим доступа: <http://ej.kubagro.ru/2012/09/pdf/25.pdf>, 1,812 у.п.л.

12. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №04(088). С. 340 – 359. – IDA [article ID]: 0881304022. – Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.

13. 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. – Режим доступа: <http://elibrary.ru/item.asp?id=18632909>.

14. Трошин Л.П. Три сибса современного частного виноградарства России и Украины / Л.П. Трошин, А.В. Милованов, Б.А. Маховицкий // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №05(089). С. 463 – 482. – IDA [article ID]: 0891305032. – Режим доступа: <http://ej.kubagro.ru/2013/05/pdf/32.pdf>, 1,25 у.п.л., импакт-фактор РИНЦ=0,581.

15. Трошин Л. П. Интерактивная ампелография – наука и педагогика // [Интерактивная](#) ампелография и селекция винограда. – Краснодар: КубГАУ, 2012. – С. 215-221.

16. Трошин Л.П. Модернизация столового сортимента для фермерского и приусадебного виноградарства: перспективные сорта-генеты Кострикина-Крайнова / Трошин Л.П. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – № 08 (102). – IDA [article ID]: 1021408036. – Режим доступа: <http://ej.kubagro.ru/2014/08/pdf/36.pdf>, 2,188 у.п.л.

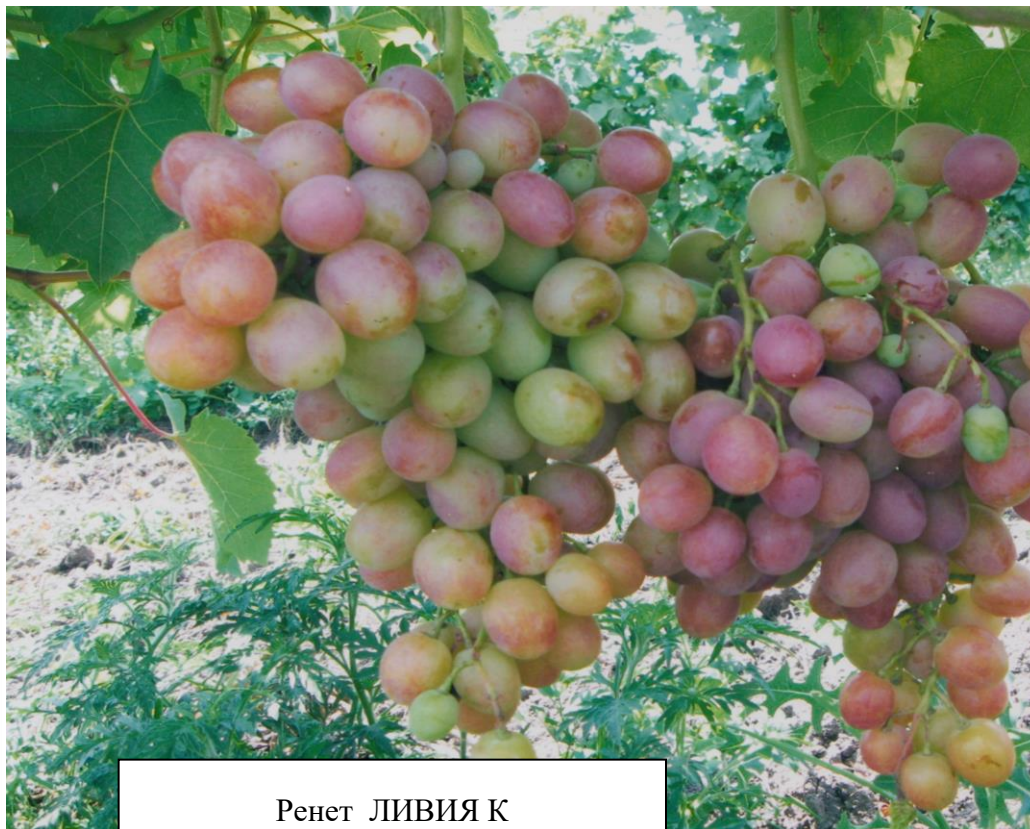
17. Трошин Л.П. Модернизация столового сортимента для фермерского и приусадебного виноградарства: перспективные сорта-генеты Кострикина-Павловского / Трошин Л.П. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – № 08 (102). – IDA [article ID]: 1021408037. – Режим доступа: <http://ej.kubagro.ru/2014/08/pdf/37.pdf>, 1,688 у.п.л.

18. Юнг К.Г. Человек и его символы. М., 1997.

19. Луценко Е.В. Лабораторный практикум по интеллектуальным информационным системам: Учебное пособие для студентов специальности "Прикладная информатика (по областям)" и другим экономическим специальностям. 2-е изд., перераб. и доп. – Краснодар: КубГАУ, 2006. – 318с. – Режим доступа: <http://elibrary.ru/item.asp?id=21683721>

20. Луценко Е.В. Когнитивные функции как обобщение классического понятия функциональной зависимости на основе теории информации в системной нечеткой интервальной математике / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №01(095). С. 122 – 183. – IDA [article ID]: 0951401007. – Режим доступа: <http://ej.kubagro.ru/2014/01/pdf/07.pdf>, 3,875 у.п.л.

21. Луценко Е.В. Решение задач ампелографии с применением АСК-анализа изображений листьев по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко, Д.К. Бандык, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №08(112). С. 862 – 910. – IDA [article ID]: 1121508064. – Режим доступа: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>, 3,062 у.п.л.



Ренет ЛИВИЯ К

3. КОЛИЧЕСТВЕННОЕ ИЗМЕРЕНИЕ СХОДСТВА-РАЗЛИЧИЯ КЛОНОВ ВИНОГРАДА ПО КОНТУРАМ ЛИСТЬЕВ С ПРИМЕНЕНИЕМ АСК-АНАЛИЗА И СИСТЕМЫ «ЭЙДОС»

В главе рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и реализующего их программного инструментария – интеллектуальной системы «Эйдос» для решения одной из важных задач ампелографии: количественного определения сходства-различия различных клонов винограда по контурам листьев. Для решения этой задачи выполняются следующие этапы:

1) оцифровка сканированных изображений листьев и создание их математических моделей;

2) формирование математических моделей конкретных листьев с применением теории информации;

3) формирование моделей обобщенных образов листьев различных клонов на основе конкретных листьев (многопараметрическая типизация);

4) верификация модели путем идентификации конкретных листьев с обобщенными образами клонов, т.е. классами (системная идентификация);

5) количественное определение сходства-различия клонов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных клонов. Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о клоне, к которому он относится, включающее как информацию об истинной форме листа данного клона (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным воздействием окружающей среды. Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение сигнала об истинной форме листа каждого клона на основе ряда

зашумленных конкретных примеров листьев данного клона. Таким образом, создается один образ формы листа каждого клона, независящий от их конкретных реализаций, т.е. «Эйдос» этих изображений (в смысле Платона) - прототип или архетип (в смысле Юнга) изображений [27].

3.1. Формулировка проблемы

В ампелографии существует проблема количественного определения сходства-различия различных клонов винограда.

Обычно сравнение клонов винограда по контурам листьев проводят путем кластерного анализа конкретных реализаций этих клонов, т.е. самих листьев. При этом из-за невнимательности или намеренно забывают о том, что форма листа зависит не только от клона, но и от множества различных внутренних и внешних факторов. Это явление, из-за которого даже на одном кусте нет двух одинаковых листьев, хорошо известно и называется «полиморфизм». Поэтому при применении кластерного анализа листьев на его результаты будут оказывать все эти факторы, а не только то, к каким клонам относятся эти листья [20, 26].

Существуют различные варианты выбора исходных данных о клонах для решения этой проблемы, не понятно только, с помощью какого измерительного (биометрического) инструментария решать эту проблему.

3.2. Предпосылки и идея решения проблемы

Предпосылки

В работе [1] подробно рассматривается применение АСК-анализа и системы «Эйдос» для синтеза и применения адаптивных интеллектуальных измерительных систем в различных предметных областях, т.е. для многопараметрической типизации и системной идентификации сложных систем. Предлагается применить эту интеллектуальную измерительную технологию для решения поставленной проблемы ампелографии.

В работе [2] предлагается *новый подход* к оцифровке изображений листьев, основанный на использовании полярной

системы координат, центра тяжести изображения и его внешнего контура и подробно описывается *технология* применения АСК-анализа, его математической модели – системной теории информации и программного инструментария – интеллектуальной системы «Эйдос» для решения ряда важных задач ампелографии: 1) оцифровка сканированных изображений листьев и создание их математических моделей; 2) формирование математических моделей конкретных листьев с применением теории информации; 3) формирование моделей обобщенных образов листьев различных клонов; 4) сравнение образа конкретного листа с обобщенным образом листа разных клонов и определение количественной степени сходства-различия между ними, т.е. идентификация клона по листу; 5) количественное определение сходства-различия клонов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных клонов сорта.

Поэтому в данной работе мы не будем описывать саму технологию применения АСК-анализа и системы «Эйдос», т.к. они подробно описаны в статье [1] и других работах авторов⁸, а больше внимания уделим описанию самого ампелографического смысла сформулированной проблемы и ее решения.

Идея

Возникает идея сравнивать клоны не по конкретным листьям, а по обобщенным образам клонов, сформированным путем многопараметрической типизации. В этих обобщенных образах влияние полиморфизма будет нивелировано и отфильтровано как шум, а все специфическое для исследуемых клонов выдвинуто на первый план. Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о клоне, к которому он относится, включающее как информацию об истинной форме листа данного клона (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным

⁸ См., например: <http://lc.kubagro.ru/> <http://lc.kubagro.ru/aidos/Aidos-X.htm>

воздействием окружающей среды. В разделе рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и реализующего их программного инструментария – интеллектуальной системы «Эйдос» для решения одной из важных задач ампелографии: количественного определения сходства-различия различных клонов винограда по контурам листьев.

В качестве исходных данных о клонах винограда для решения поставленной проблемы предлагается использовать сканированные изображения листьев, а в качестве инструментария для построения измерительной системы и ее применения – АСК-анализ и систему «Эйдос».

Другими вариантами исходных данных могут быть геном или такие фенотипические признаки, как форма куста, форма и размер грозди, форма, цвет и размер ягод и другие. Однако контуры листьев выбраны потому, что, с одной стороны, являются во многих отношениях наиболее удобным объектом исследования, а, с другой стороны, содержат информацию и о геноме, а значит косвенно и о других фенотипических признаках, обусловленных как геномом, так и воздействием окружающей среды.

3.3. Исходные данные

Общеизвестно, что в клоновой селекции винограда – ныне основным фактором повышения продуктивности виноградных насаждений – до сегодняшнего дня центральной ее задачей остается отбор плюс-трангрессивных высокопродуктивных растений, так называемых протоклонов, с целью дальнейшего их размножения и внедрения в производство. При этом в вегетативном потомстве размножаемых протоклонов из-за высокой модификационной изменчивости количественных признаков и наличия высокой степени полиморфизма листьев возникает вопрос об их различимости между собой по комплексу

морфологических признаков листьев – самых распространенных определителей самих генотипов [8-19].

Цель исследования: установить схожесть или степень различия по комплексу морфологических признаков листьев вегетативного потомства плюс-трансгрессивных растений четырех протоклонов сорта Каберне-Совиньон (110, 177, 221 и 838).

Материалом исследований служили листья первого вегетативного потомства вышеназванных четырех протоклонов, окончательно отобранных селекционером СКЗНИИСиВ Т.А.Нудьга в 2007 г. Взрослые листья вегетативного потомства протоклонов были заготовлены нами осенью 2015 года, сканированы, измерены с помощью компьютерной программы SIAMS⁹ и в последующем проанализированы методом АСК-анализа [3, 4] с применением его программного инструментария – интеллектуальной системы «Эйдос» [5] (рис. 1-4).

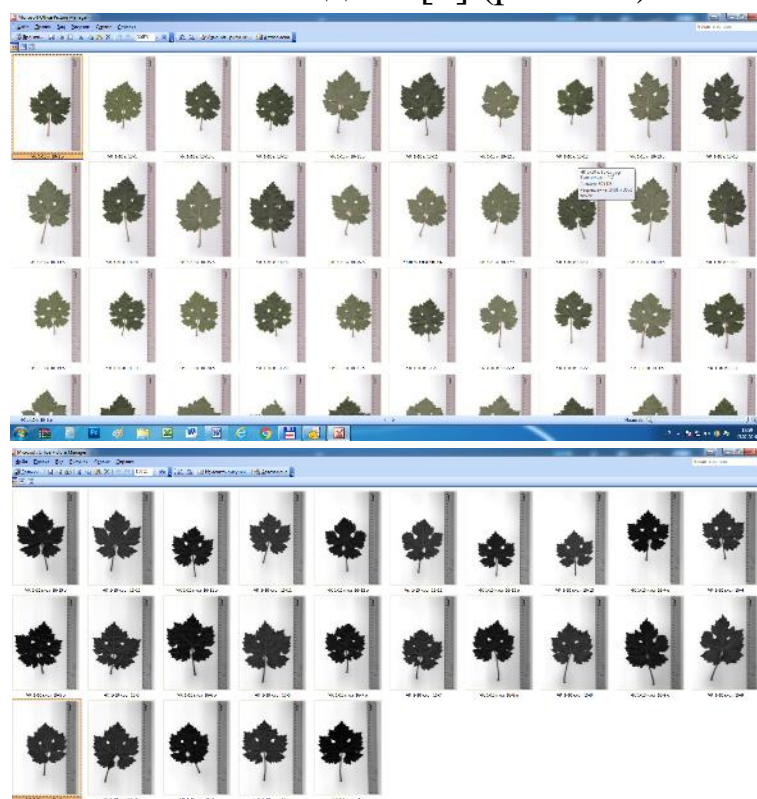


Рис. 1. Листья клона № 110 кустов 10-1 (вверху) и 16-2 (внизу)

⁹ <http://new.siams.com/>



Рис. 2. Листья клона № 177 кустов 7 (вверху) и 9 (внизу)

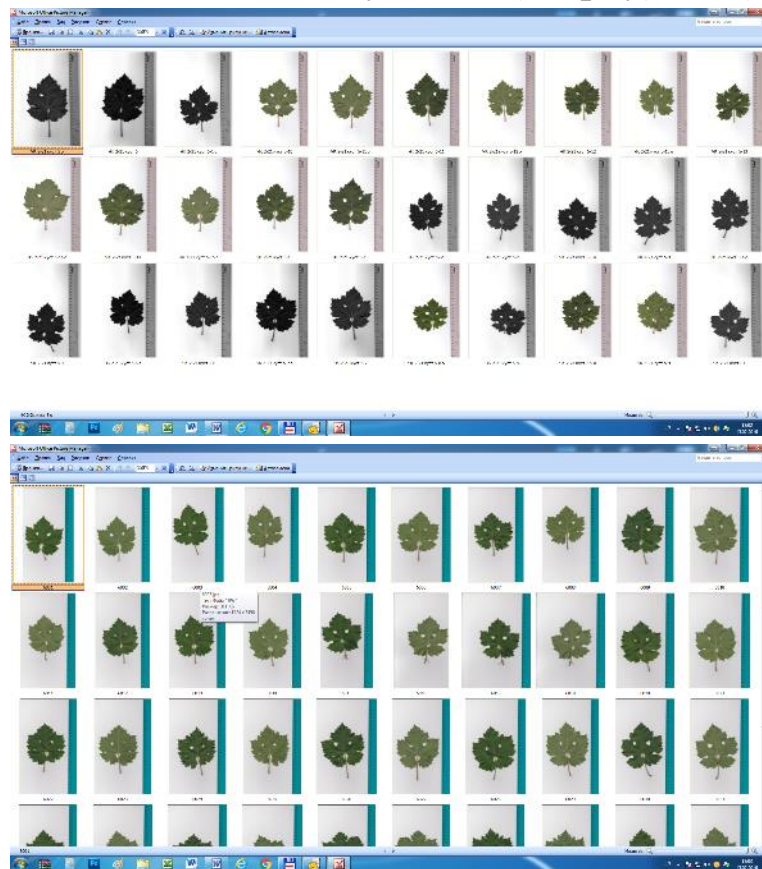


Рис. 3. Листья клона № 221 кустов 5 (вверху) и 6 (внизу)

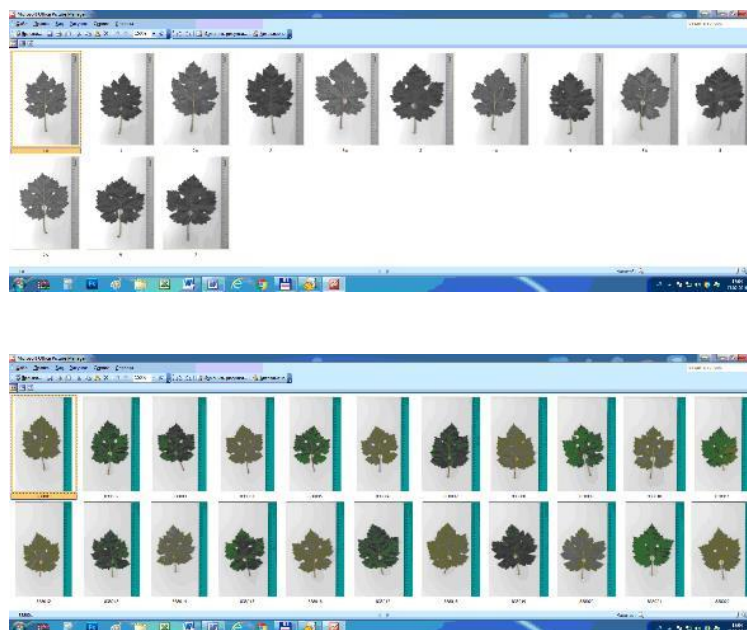


Рис. 4. Листья клона № 838 кустов 1 (вверху) и 6 (внизу)

3.4. Решение проблемы

Решение проблемы, сформулированной в разделе, осуществлено в три итерации.

На первой итерации создана исходная модель, включающая 8 обобщенных образов классов по северной и южной сторонам кустов, конкретные реализации которых в виде отдельных листьев приведены на рисунках 1 - 4.

На второй итерации создана модель, включающая 4 обобщенных образа класса по кустам, конкретные реализации которых в виде отдельных приведены на рисунках 1 - 4.

На третьей итерации создана итоговая модель, включающая 3 обобщенных образа класса по клонам, выявленным на второй итерации.

3.4.1. Первая итерация

Этапы 1-3 решения проблемы: 1) оцифровка сканированных изображений листьев и создание их математических моделей; 2) формирование математических моделей конкретных листьев с применением теории информации мы в данной работе рассматривать не будем, т.к. они аналогичны подробно описанным в статье [2].

Отметим лишь, что на этих этапах были автоматически созданы классификационные и описательные шкалы и градации, в наглядной форме приведенные на рис. 5.

На рисунке 5 цветом обозначена ценность различных градаций описательных шкал для того, чтобы отличить один класс от другого. На верхнем рисунке приведены все градации всех описательных шкал, а на нижнем – только наиболее ценные в каждой шкале.

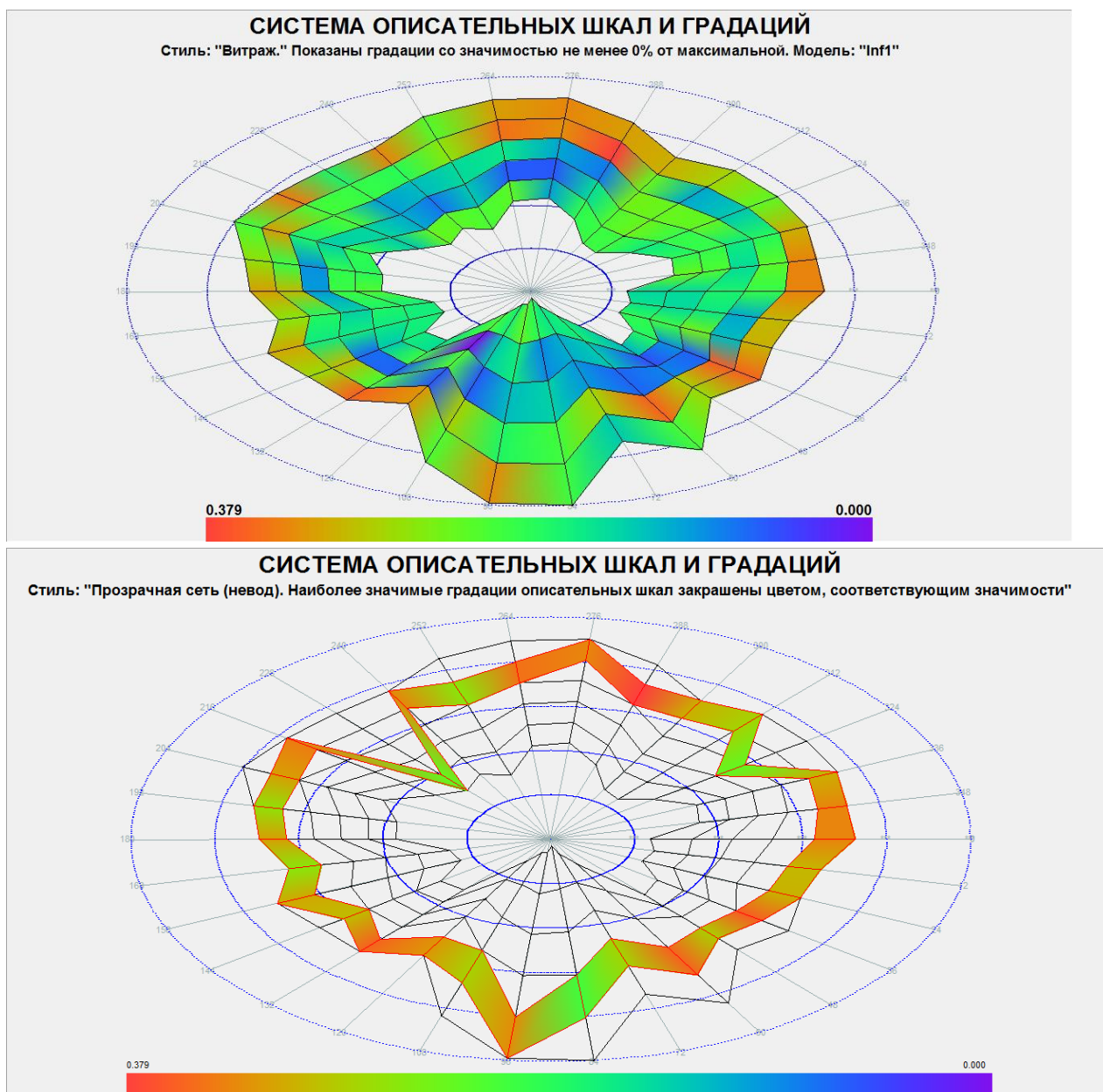


Рис. 5. Классификационные и описательные шкалы и градации при 8 классах

Этап 3) формирование моделей обобщенных образов листьев различных клонов на основе конкретных листьев (многопараметрическая типизация).

В результате многопараметрической типизации были получены следующие обобщенные образы листьев клонов по северным (вверху) и южным (внизу) кустам (рис. 6 - 9).

На рисунках 6 - 9 цветом обозначена ценность различных градаций описательных шкал для того, чтобы отличить данный класс от остальных. На правом рисунке приведены все градации всех описательных шкал, а на нижнем – только наиболее характерные для данного класса в каждой шкале, т.е. наиболее ценные в каждой шкале для того чтобы отличить данный класс от остальных.

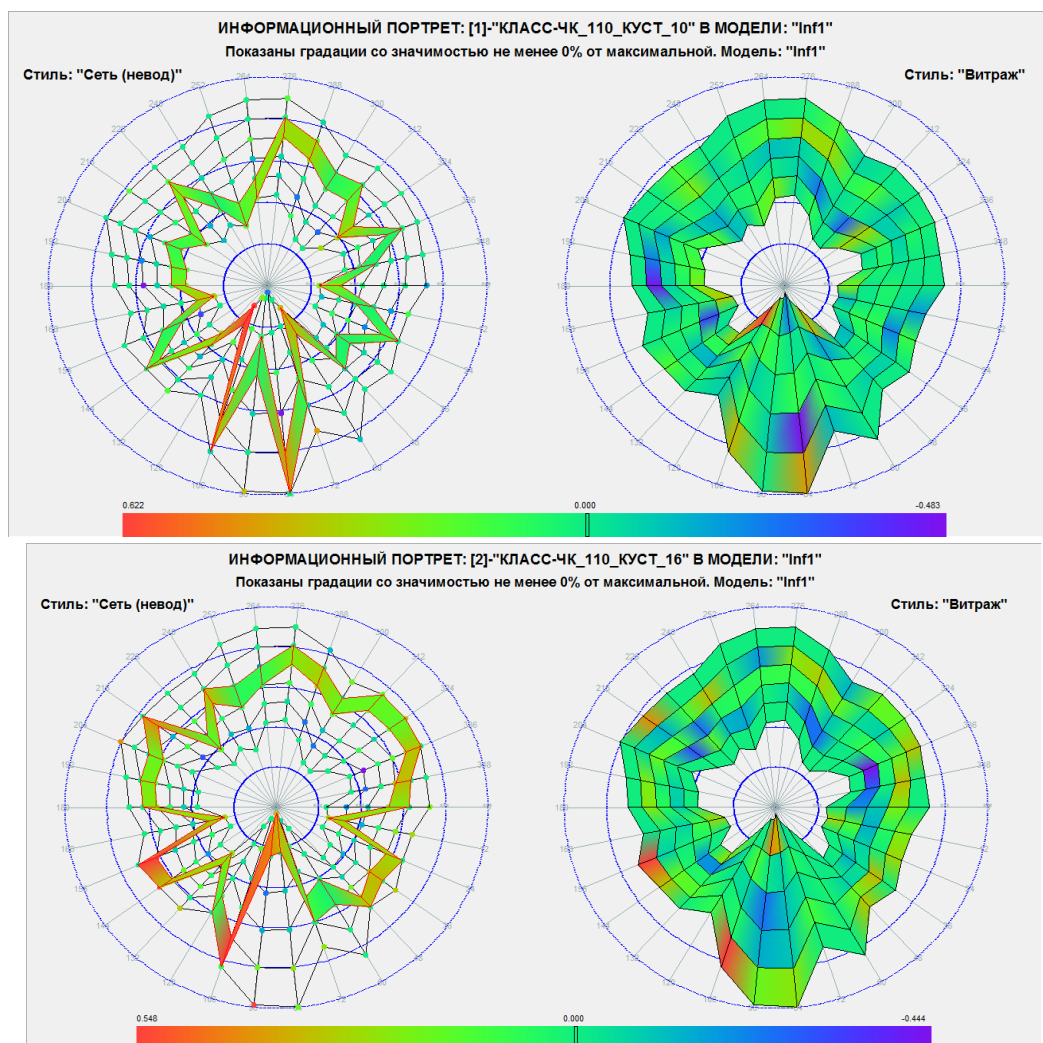


Рис. 6. Обобщенные образы листьев клона 110 по северным (вверху) и южным (внизу) кустам

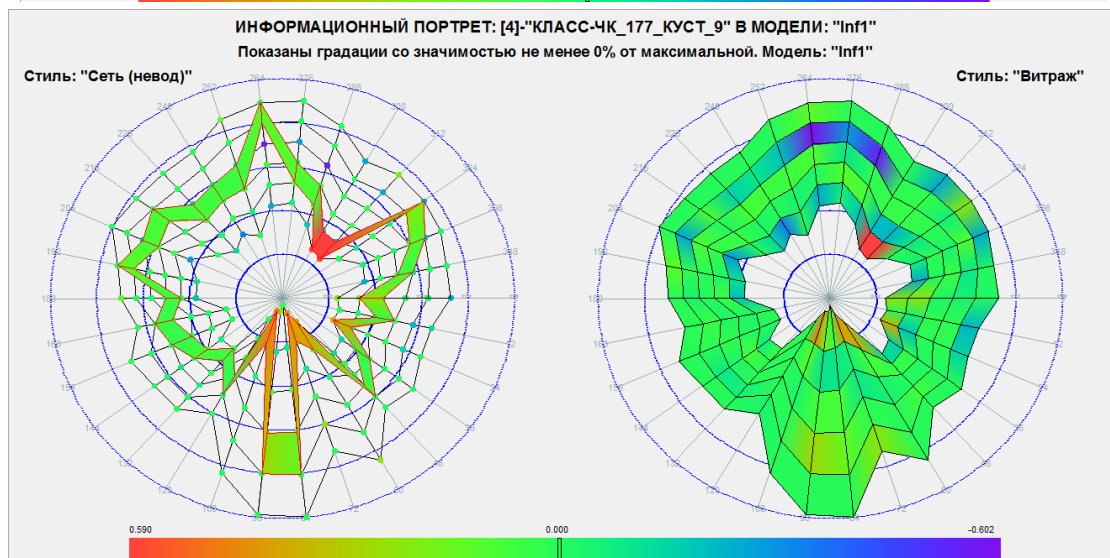
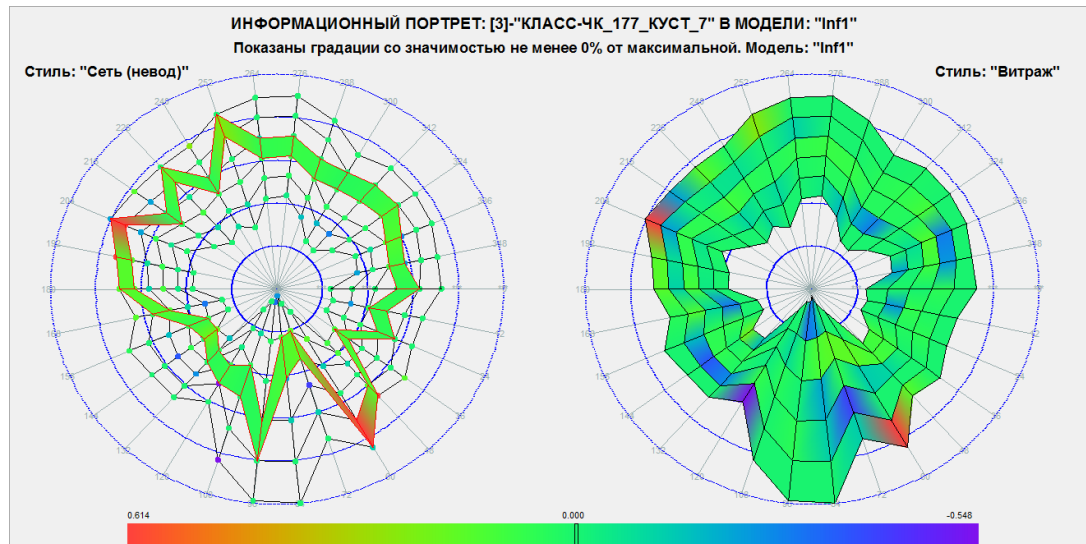
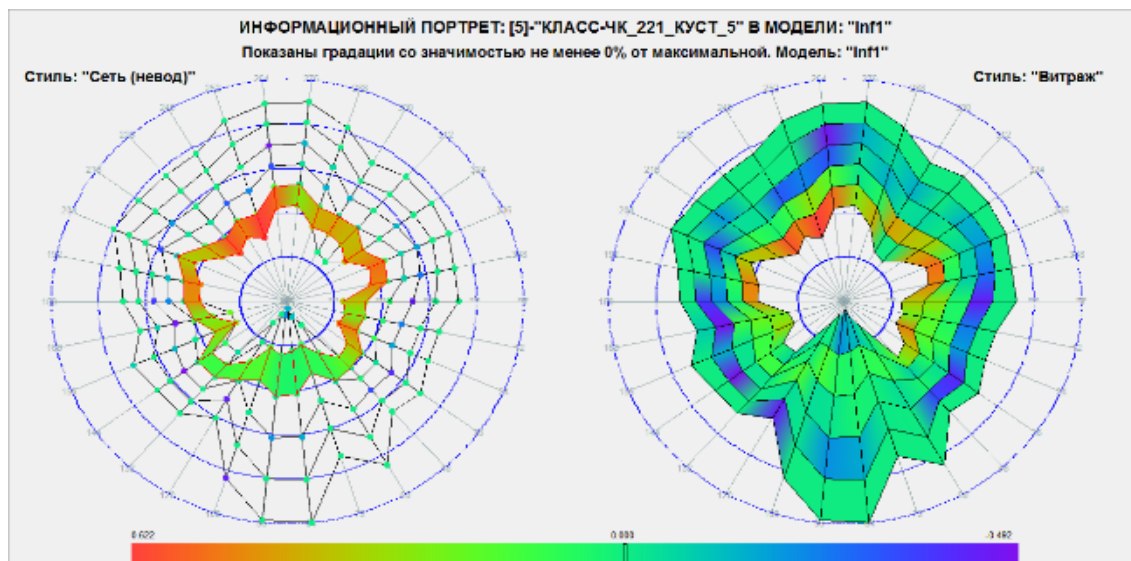


Рис. 7. Обобщенные образы листьев клона 117 по северным (вверху) и южным (внизу) кустам



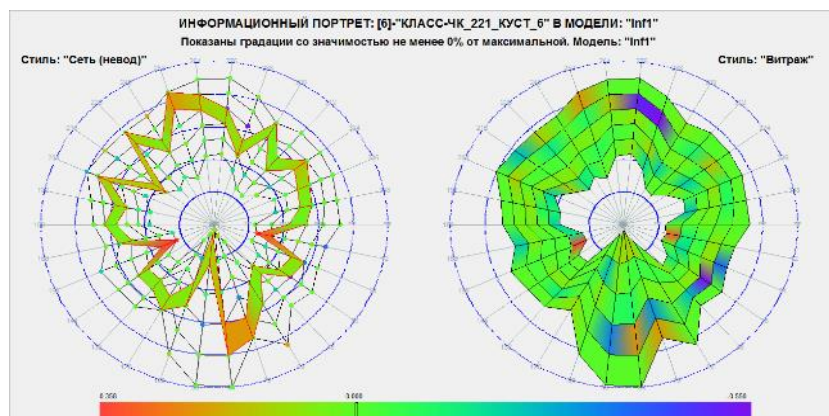


Рис. 8. Обобщенные образы листьев клона 221 по северным (вверху) и южным (внизу) кустам

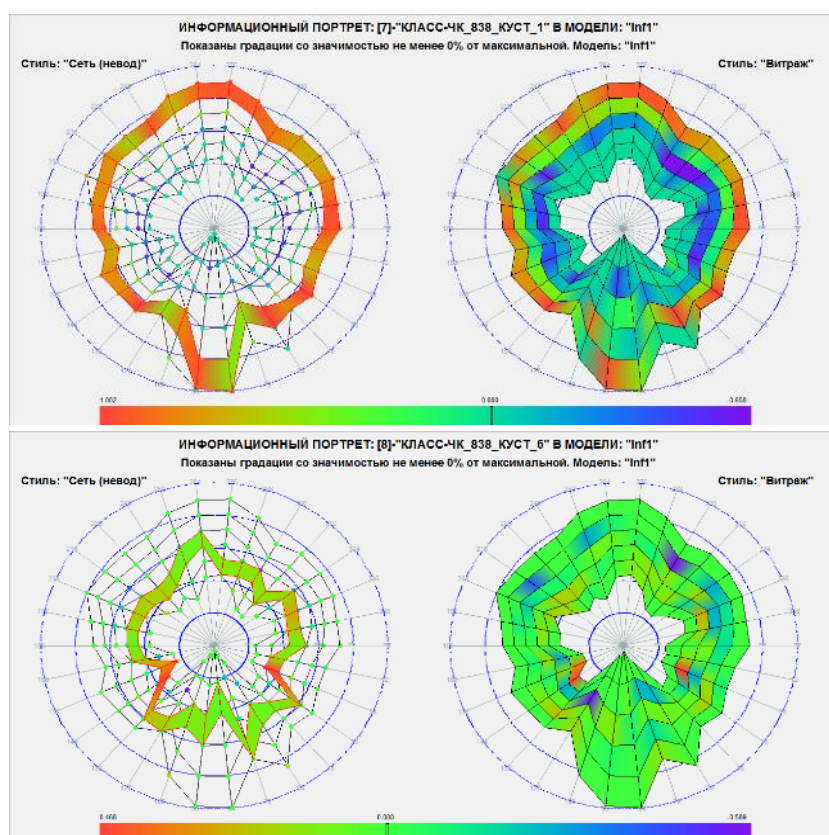


Рис. 9. Обобщенные образы листьев клона 838 по северным (вверху) и южным (внизу) кустам

Этап 4) верификация модели путем идентификации конкретных листьев с обобщенными образами клонов, т.е. классами (системная идентификация).

Созданная модель показала достаточно высокую достоверность, около 70% несмотря на очень небольшое

количество измерительных шкал (всего 30) и градаций (5) (рис 10):

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентификац... | Вероятность правильной не идентиф... | Средняя вероятно... результата | Дата получения результата | Время получения результ... |
|---|----------------------------------|---------------------------------------|--------------------------------------|--------------------------------|---------------------------|----------------------------|
| ABS - частный критерий: количество ветвей соцветий: "Клас... | Корреляция абс частот с обр... | 100,000 | 5,540 | 52,770 | 13.02.2016 | 09:59:16 |
| 1. ABS - частный критерий: количество ветвей соцветий: "Клас... | Сумма абс частот по признак... | 100,000 | | 50,000 | 13.02.2016 | 09:59:16 |
| 2. PFC1 - частный критерий: усл. вероятность i-го признака сред... | Корреляция усл.отн частот с о... | 100,000 | 5,540 | 52,770 | 13.02.2016 | 09:59:26 |
| 2. PFC1 - частный критерий: усл. вероятность i-го признака сред... | Сумма усл.отн частот по при... | 100,000 | | 50,000 | 13.02.2016 | 09:59:26 |
| 3. PFC2 - частный критерий: условная вероятность i-го признака... | Корреляция усл.отн частот с о... | 100,000 | 5,540 | 52,770 | 13.02.2016 | 09:59:36 |
| 3. PFC2 - частный критерий: условная вероятность i-го признака... | Сумма усл.отн частот по при... | 100,000 | | 50,000 | 13.02.2016 | 09:59:36 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу: в... | Семантический резонанс: зна... | 80,469 | 61,191 | 70,830 | 13.02.2016 | 09:59:42 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу: в... | Сумма знаний | 80,469 | 64,406 | 72,438 | 13.02.2016 | 09:59:42 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу: в... | Семантический резонанс: зна... | 80,469 | 61,110 | 70,789 | 13.02.2016 | 09:59:52 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу: в... | Сумма знаний | 80,859 | 64,280 | 72,570 | 13.02.2016 | 09:59:52 |
| 6. INF3 - частный критерий: Хи-квадрат: разности между фактик... | Семантический резонанс: зна... | 90,625 | 43,986 | 67,306 | 13.02.2016 | 09:59:58 |
| 6. INF3 - частный критерий: Хи-квадрат: разности между фактик... | Сумма знаний | 90,625 | 43,986 | 67,306 | 13.02.2016 | 09:59:58 |
| 7. INF4 - частный критерий: ROI (Return On Investment): веротно... | Семантический резонанс: зна... | 60,547 | 79,221 | 69,884 | 13.02.2016 | 10:00:05 |
| 7. INF4 - частный критерий: ROI (Return On Investment): веротно... | Сумма знаний | 91,016 | 46,783 | 68,899 | 13.02.2016 | 10:00:05 |
| 8. INF5 - частный критерий: ROI (Return On Investment): веротно... | Семантический резонанс: зна... | 60,547 | 79,221 | 69,884 | 13.02.2016 | 10:00:16 |
| 8. INF5 - частный критерий: ROI (Return On Investment): веротно... | Сумма знаний | 91,016 | 46,803 | 68,910 | 13.02.2016 | 10:00:16 |
| 9. INF6 - частный критерий: разн усл и без усл вероятностей; вер... | Семантический резонанс: зна... | 85,547 | 48,912 | 67,229 | 13.02.2016 | 10:00:24 |
| 9. INF6 - частный критерий: разн усл и без усл вероятностей; вер... | Сумма знаний | 90,625 | 40,901 | 65,763 | 13.02.2016 | 10:00:24 |
| 10. INF7 - частный критерий: разн усл и без усл вероятностей; ве... | Семантический резонанс: зна... | 85,547 | 48,831 | 67,189 | 13.02.2016 | 10:00:30 |
| 10. INF7 - частный критерий: разн усл и без усл вероятностей; ве... | Сумма знаний | 90,625 | 40,700 | 65,662 | 13.02.2016 | 10:00:30 |

Рис. 10. Достоверность созданных моделей 1-й итерации

Этап 5) количественное определение сходства-различия клонов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных клонов (рис. 11).

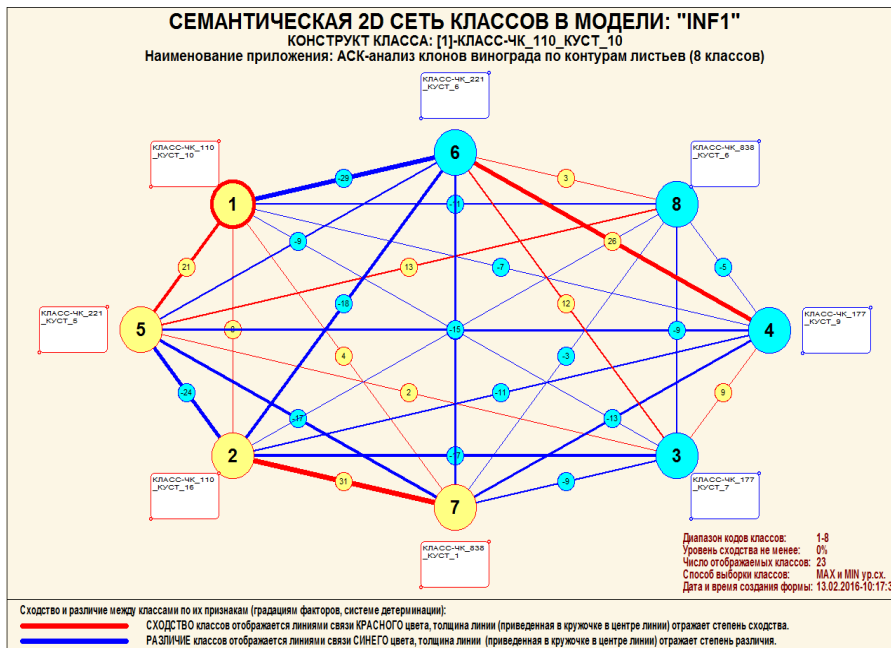


Рис. 11. Результаты кластерно-конструктивного анализа по 1-й модели

Из анализа формы на рис. 11 мы видим, что:

- 1) форма листьев одного клона с кустов северных и южных склонов сильно отличается;
- 2) форма листьев с кустов северных и южных склонов у разных клонов имеет много общего.

Из этого результата можно сделать вывод о том, что на форму контура листьев клон оказывает меньшее влияние, чем расположение влияния окружающей среды в виде освещенности, связанной с расположением куста на северном или южном склоне холма.

Основываясь на этом результате с целью решения сформулированной проблемы было принято решение о создании 2-й модели, в которой многопараметрическая типизация была бы проведена по кустам независимо от их расположения и влияние бы этого расположения и связанной с ним освещенности было бы нивелировано, т.е. отфильтровано как шум.

Отметим, что если бы в исследовании ставилась задача исследовать влияние освещенности на форму листа, то при создании модели была бы выбрана другая группировка листьев по классам.

3.4.2. Вторая итерация

Этапы 1-3 решения проблемы:

- 1) оцифровка сканированных изображений листьев и создание их математических моделей;
- 2) формирование математических моделей конкретных листьев с применением теории информации мы в данной работе рассматривать не будем, т.к. они аналогичны подробно описанным в статье [2].

Отметим лишь, что на этих этапах были автоматически созданы классификационные и описательные шкалы и градации, в наглядной форме приведенные на рис. 12:

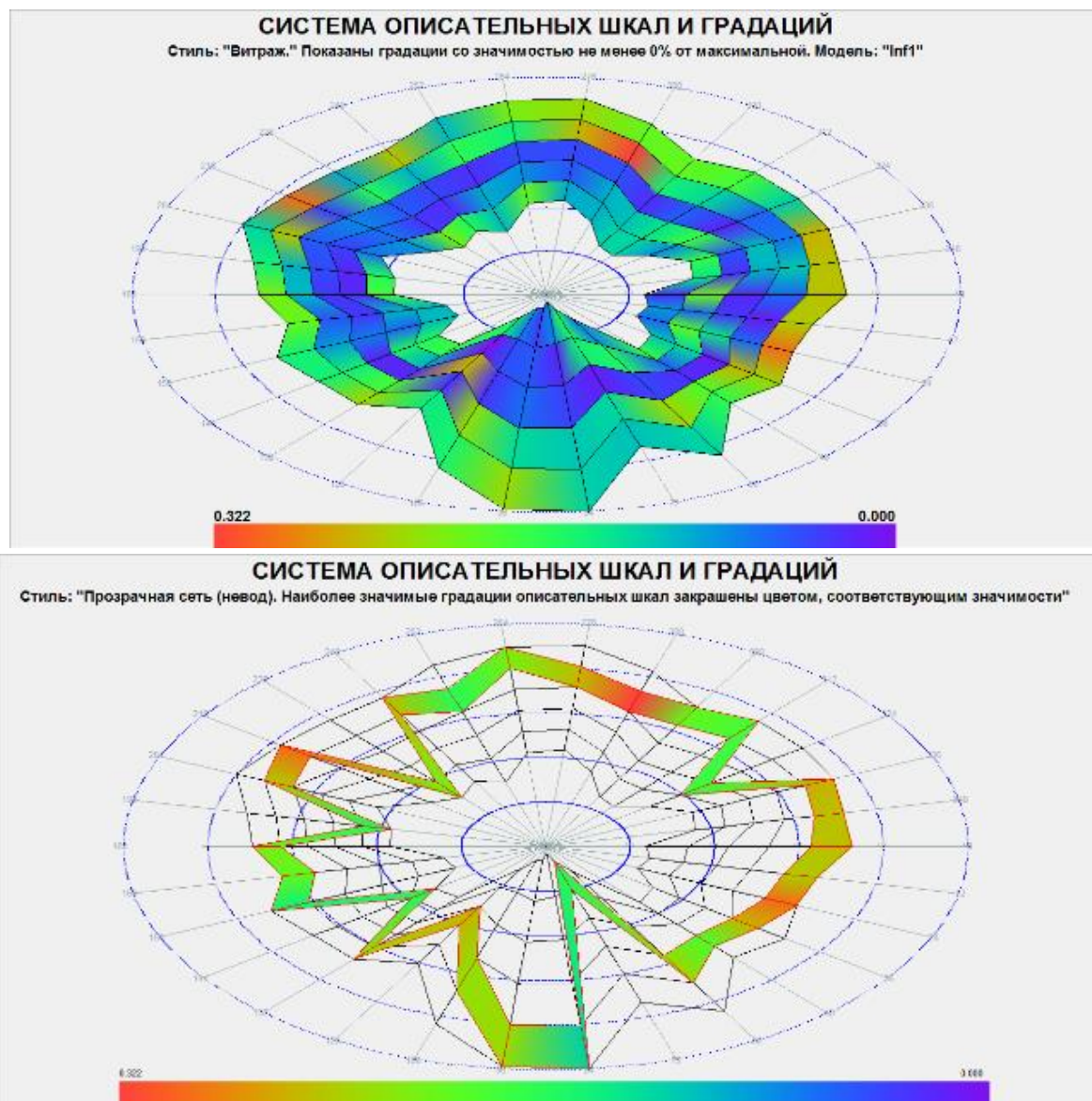


Рис. 12. Классификационные и описательные шкалы и градации при 4 классах

Этап 3) формирование моделей обобщенных образов листьев различных клонов на основе конкретных листьев (многопараметрическая типизация).

В результате многопараметрической типизации были получены следующие обобщенные образы листьев клонов (рис. 13 – 16):

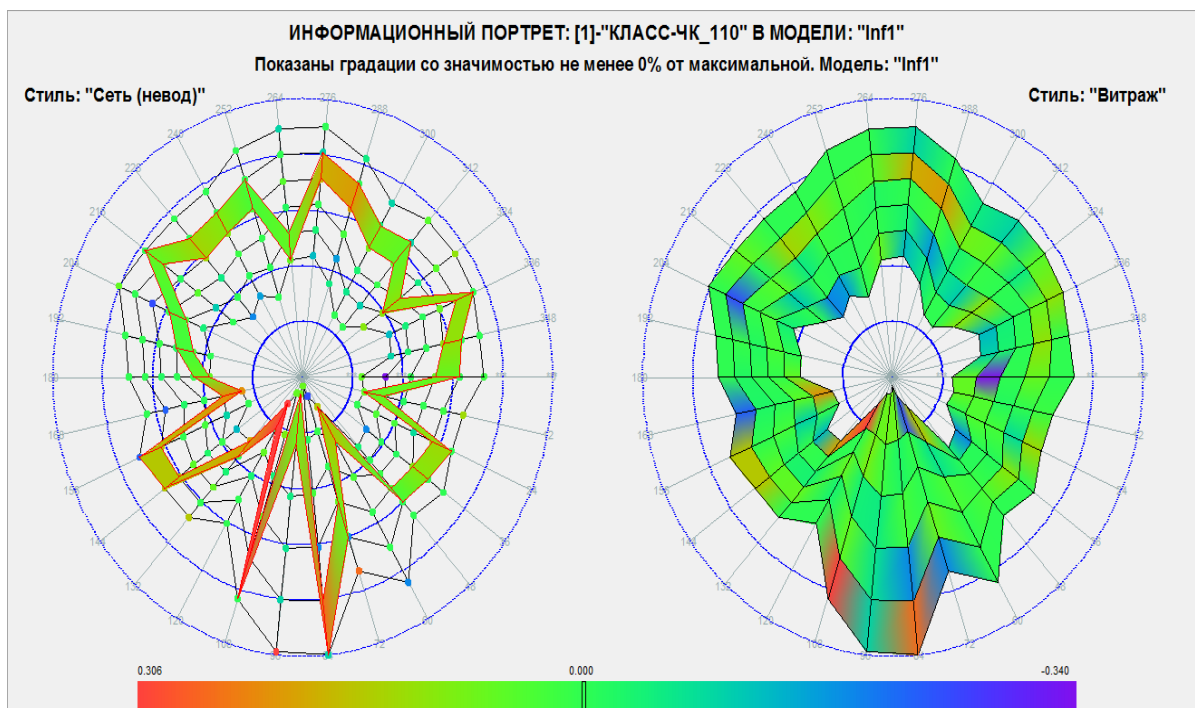


Рис. 13. Обобщенные образы листьев клона 110

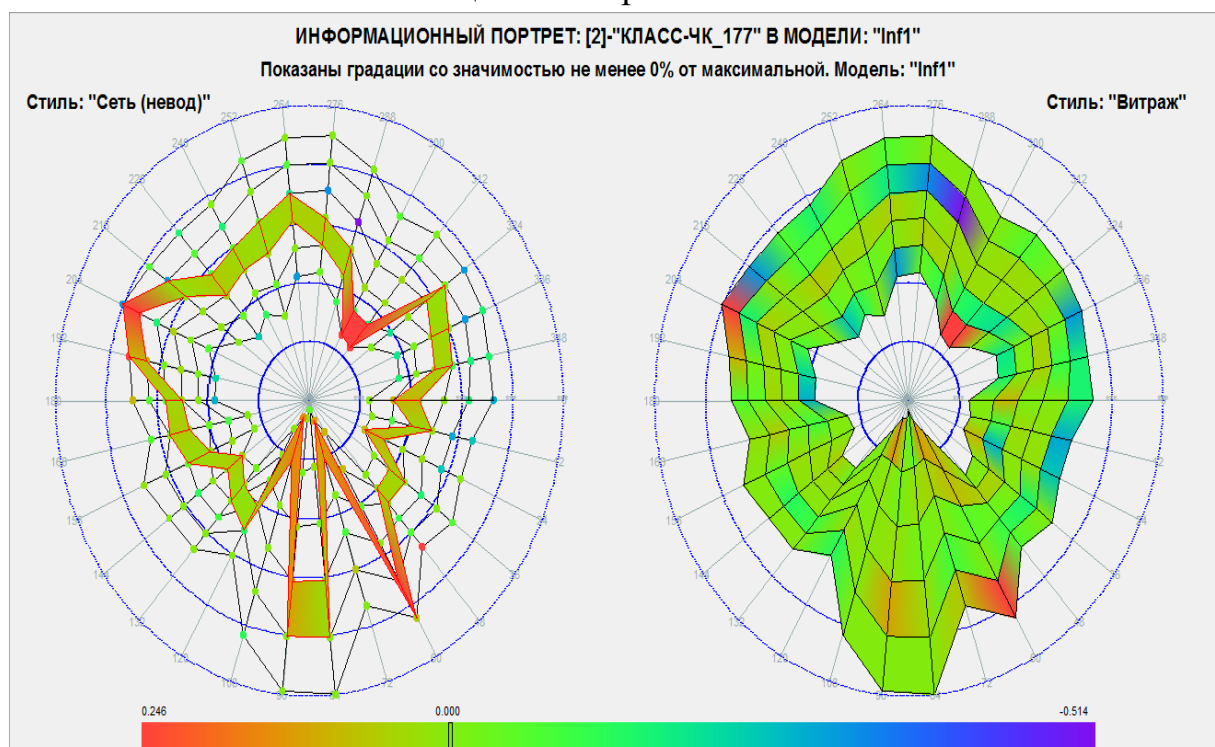


Рис. 14. Обобщенные образы листьев клона 117

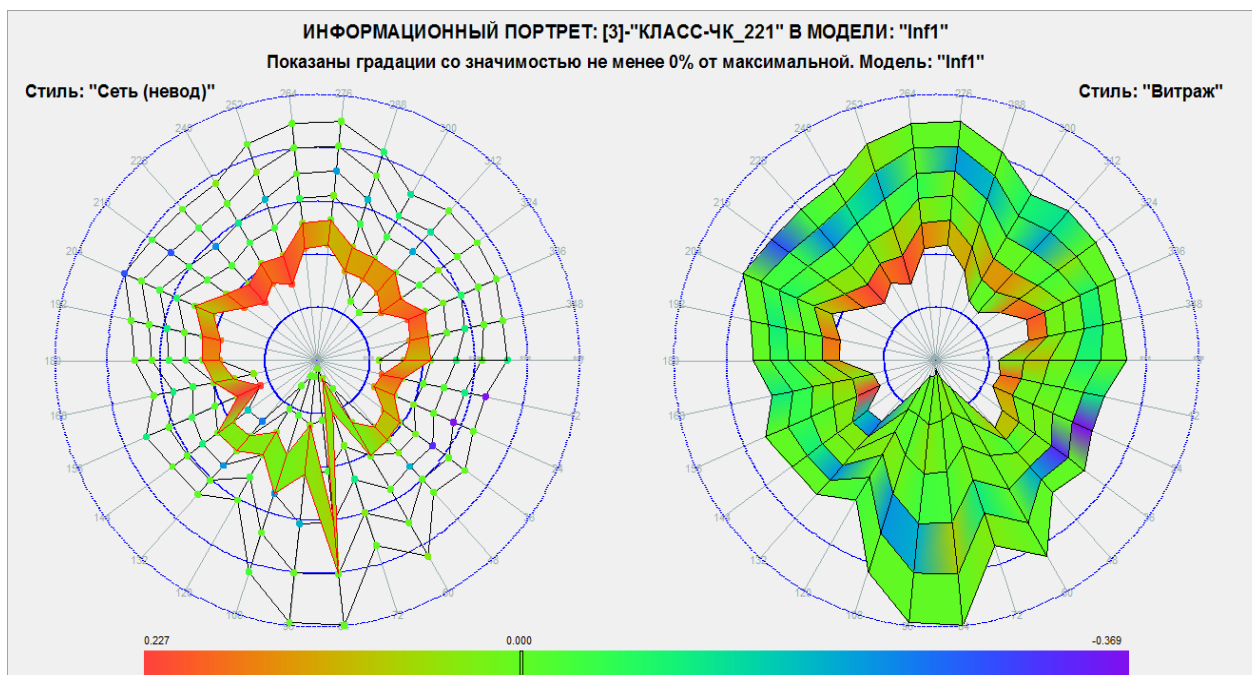


Рис. 15. Обобщенные образы листьев клона 221

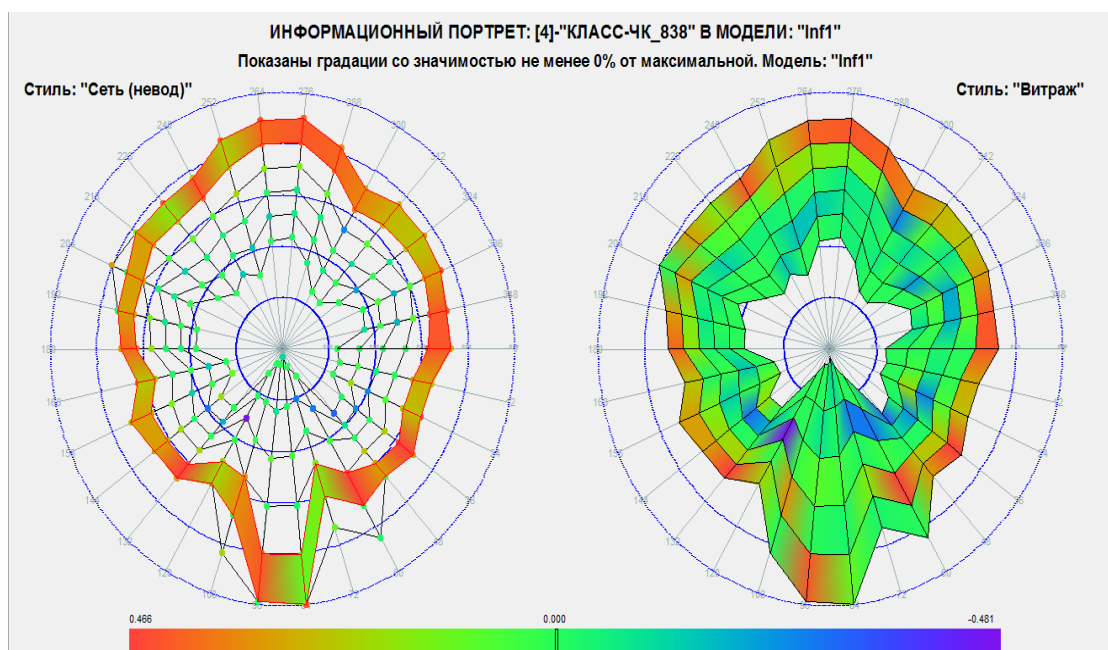


Рис. 16. Обобщенные образы листьев клона 838

Этап 4) верификация модели путем идентификации конкретных листьев с обобщенными образами клонов, т.е. классами (системная идентификация).

Созданная 2-я модель показала достаточно высокую достоверность, близкую к 1-й модели: около 70% несмотря на очень небольшое количество измерительных шкал (всего 30) и градаций (5) (рис 17):

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентификации... | Вероятность правильной не идентиф... | Средняя вероятность признан... результата | Дата получения результата | Время получения результ... |
|--|----------------------------------|---|--------------------------------------|---|---------------------------|----------------------------|
| ABS - частный критерий: количество встреч сочетаний "клас... | Корреляция абс частот с обр... | 99.609 | 3.165 | 51.387 | 13.02.2016 | 10:14:41 |
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Сумма абс частот по признак... | 100.000 | | 50.000 | 13.02.2016 | 10:14:41 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Корреляция усл отн частот с о... | 99.609 | 3.165 | 51.387 | 13.02.2016 | 10:14:46 |
| 2. PRC1 - частный критерий: усл. вероятность i-го признака сред... | Сумма усл отн частот по при... | 100.000 | | 50.000 | 13.02.2016 | 10:14:46 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Корреляция усл отн частот с о... | 99.609 | 3.165 | 51.387 | 13.02.2016 | 10:14:52 |
| 3. PRC2 - частный критерий: условная вероятность i-го признака... | Сумма усл отн частот по при... | 100.000 | | 50.000 | 13.02.2016 | 10:14:52 |
| 4. INF1 - частный критерий: количество знаний по А.Жаркевичу; в... | Семантический резонанс зна... | 71.875 | 55.926 | 63.901 | 13.02.2016 | 10:14:57 |
| 4. INF1 - частный критерий: количество знаний по А.Жаркевичу; в... | Сумма знаний | 70.703 | 66.175 | 68.439 | 13.02.2016 | 10:14:57 |
| 5. INF2 - частный критерий: количество знаний по А.Жаркевичу; в... | Семантический резонанс зна... | 71.875 | 55.926 | 63.901 | 13.02.2016 | 10:15:03 |
| 5. INF2 - частный критерий: количество знаний по А.Жаркевичу; в... | Сумма знаний | 70.313 | 65.822 | 68.067 | 13.02.2016 | 10:15:03 |
| 6. INF3 - частный критерий: Хинквадрат, разности между факти... | Семантический резонанс зна... | 75.391 | 51.020 | 63.205 | 13.02.2016 | 10:15:09 |
| 6. INF3 - частный критерий: Хинквадрат, разности между факти... | Сумма знаний | 75.391 | 51.020 | 63.205 | 13.02.2016 | 10:15:10 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятн... | Семантический резонанс зна... | 59.375 | 72.671 | 66.023 | 13.02.2016 | 10:15:15 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятн... | Сумма знаний | 77.344 | 57.163 | 67.253 | 13.02.2016 | 10:15:15 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятн... | Семантический резонанс зна... | 59.375 | 72.513 | 65.944 | 13.02.2016 | 10:15:20 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятн... | Сумма знаний | 77.344 | 57.163 | 67.253 | 13.02.2016 | 10:15:20 |
| 9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер... | Семантический резонанс зна... | 72.656 | 56.929 | 64.792 | 13.02.2016 | 10:15:26 |
| 9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер... | Сумма знаний | 75.391 | 50.188 | 62.789 | 13.02.2016 | 10:15:26 |
| 10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве... | Семантический резонанс зна... | 72.266 | 56.984 | 64.625 | 13.02.2016 | 10:15:32 |
| 10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве... | Сумма знаний | 75.391 | 50.085 | 62.738 | 13.02.2016 | 10:15:32 |

Рис. 17. Достоверность созданных моделей 2-й итерации

Этап 5) количественное определение сходства-различия клонов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных клонов (рис. 18):

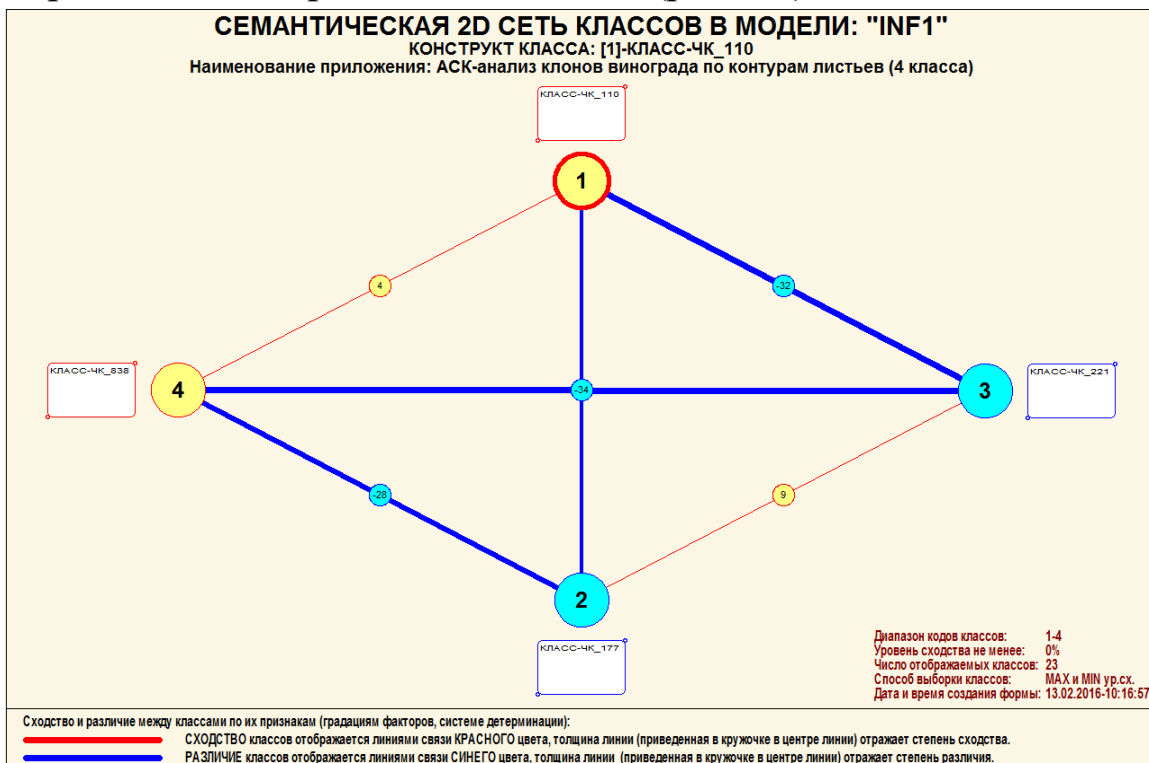


Рис. 18. Результаты кластерно-конструктивного анализа по 1-й модели

В таблице приведена матрица сходства обобщенных образов клонов (классов), отображенная в форме когнитивной диаграммы на рисунке 18.

Таблица. – Матрица сходства обобщенных образов клонов

| № | Клон | 110 | 177 | 221 | 838 |
|---|--------|--------|--------|--------|--------|
| 1 | ЧК_110 | 100 | -20,2 | -32,24 | 3,5216 |
| 2 | ЧК_177 | -20,2 | 100 | 8,7179 | -27,7 |
| 3 | ЧК_221 | -32,24 | 8,7179 | 100 | -34,22 |
| 4 | ЧК_838 | 3,5216 | -27,7 | -34,22 | 100 |

Из анализа таблицы и формы на рис. 18 мы видим, что исследованные четыре клон №№ 110, 177, 221 и 838 по комплексу признаков взрослых листьев являются заметно различными, а сходными в очень небольшой степени по этим признакам оказались лишь клоны №110 и №838, №177 и №221.

Рассматривая параллельно в течение ряда лет другой комплекс биолого-хозяйственных признаков и свойств этих же четырех клонов, селекционеры-виноградари Нудьга Т.А., Ильницкая Е.Т. и Трошин Л.П. пришли к выводу об отборе лидирующих высокопродуктивных клонов в пределах каждой популяции, которые под названиями Каберне Тамани (№ 177-36), Каберне Кубани (№ 221-36) и Каберне Черноморец (№ 838-30) были переданы на государственные испытания и в 2015 году зарегистрированы в сортовой книге Госсорткомиссии России.

Хорошее совпадение результатов выявления сходства-различия клонов винограда по форме контуров листьев и независимым методом по биолого-хозяйственным признакам и свойствам увеличивает достоверность полученных в этих исследованиях результатов.

3.5. Выводы

В главе рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и

реализующего их программного инструментария – интеллектуальной системы «Эйдос» для решения одной из важных задач ампелографии: количественного определения сходства-различия различных клонов винограда по контурам листьев. Для решения этой задачи выполняются следующие этапы:

1) оцифровка сканированных изображений листьев и создание их математических моделей;

2) формирование математических моделей конкретных листьев с применением теории информации;

3) формирование моделей обобщенных образов листьев различных клонов на основе конкретных листьев (многопараметрическая типизация);

4) верификация модели путем идентификации конкретных листьев с обобщенными образами клонов, т.е. классами (системная идентификация);

5) количественное определение сходства-различия клонов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных клонов.

Форма контура конкретного листа рассматривается как зашумленное информационное сообщение о клоне, к которому он относится, включающее как информацию об истинной форме листа данного клона (чистый сигнал), так и шум, искажающий эту истинную форму, обусловленный случайным воздействием окружающей среды.

Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение сигнала об истинной форме листа каждого клона на основе ряда зашумленных конкретных примеров листьев данного клона.

Таким образом создается один образ формы листа каждого клона, независящий от их конкретных реализаций, т.е. «Эйдос»

этих изображений (в смысле Платона) - прототип или архетип (в смысле Юнга) изображений.

Описанная в главе технология синтеза и применения интеллектуальной измерительной системы может быть применена не только в ампелографии, но и в других областях, чему может способствовать и то, что система «Эйдос» размещена в полном открытом бесплатном доступе на сайте автора по адресу: <http://lc.kubagro.ru/aidos/Aidos-X.htm>.

Материалы данной работы могут быть использованы при проведении лабораторных работ по дисциплинам, связанным с интеллектуальными технологиями, представлением знаний и системами искусственного интеллекта, а также по эконометрике, биометрии, экологии, педагогике, психологии и медицине [1, 6, 7].

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Луценко Е.В. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/01.pdf>, 3,750 у.п.л.

2. Луценко Е.В. Решение задач ампелографии с применением АСК-анализа изображений листьев по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко, Д.К. Бандык, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №08(112). С. 862 – 910. – IDA [article ID]: 1121508064. – Режим доступа: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>, 3,062 у.п.л.

3. 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. <http://elibrary.ru/item.asp?id=18632909>

4. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №04(088). С. 340 – 359. – IDA [article ID]: 0881304022. – Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.

5. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос». Монография (научное издание). – Краснодар, КубГАУ. 2014. – 600 с. ISBN 978-5-94672-830-0. <http://elibrary.ru/item.asp?id=18271217>

6. Луценко Е.В. Лабораторный практикум по интеллектуальным информационным системам: Учебное пособие для студентов специальности "Прикладная информатика (по областям)" и другим экономическим специальностям. 2-е изд., перераб. и доп. – Краснодар: КубГАУ, 2006. – 318с. – Режим доступа: <http://elibrary.ru/item.asp?id=21683721>

7. Луценко Е.В. АСК-анализ влияния экологических факторов на качество жизни населения региона / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №06(110). С. 1 – 37. – IDA [article ID]: 1101506001. – Режим доступа: <http://ej.kubagro.ru/2015/06/pdf/01.pdf>, 2,312 у.п.л.

8. Трошин Л.П. Три сибса современного частного виноградарства России и Украины / Л.П. Трошин, А.В. Милованов, Б.А. Маховицкий // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №05(089). С. 463 – 482. – IDA [article ID]: 0891305032. – Режим доступа: <http://ej.kubagro.ru/2013/05/pdf/32.pdf>, 1,25 у.п.л., импакт-фактор РИНЦ=0,581.

9. Трошин Л. П. Интерактивная ампелография – наука и педагогика // [Интерактивная](#) ампелография и селекция винограда. – Краснодар: КубГАУ, 2012. – С. 215-221.

10. Трошин Л.П. Модернизация столового сортимента для фермерского и приусадебного виноградарства: перспективные сорта-генеты Кострикина-Крайнова / Трошин Л.П. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – № 08 (102). – IDA [article ID]: 1021408036. – Режим доступа: <http://ej.kubagro.ru/2014/08/pdf/36.pdf>, 2,188 у.п.л.

11. Трошин Л.П. Модернизация столового сортимента для фермерского и приусадебного виноградарства: перспективные сорта-генеты Кострикина-Павловского / Трошин Л.П. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – № 08 (102). – IDA [article ID]: 1021408037. – Режим доступа: <http://ej.kubagro.ru/2014/08/pdf/37.pdf>, 1,688 у.п.л.

12. Трошин Л.П., Музыченко А.Б., Мисливский А.И. Новации виноградарства России. 3. Клоновая селекция винограда // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – № 10 (54). – Режим доступа: <http://ej.kubagro.ru/2009/10/pdf/08.pdf>.

13. Трошин Л.П. Морфометрический анализ листовой ампелографической информации / Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №06(70). С. 460 – 490. – Режим доступа: <http://ej.kubagro.ru/2011/06/pdf/32.pdf>, 1,938 у.п.л.

14. Трошин Л.П. Морфометрический анализ листовой ампелографической информации // Виноделие и виноградарство. – 2011. - № 3. – С. 48-49; - № 4. – С. 47-49.

15. Трошин Л.П. Leaf morphometry of kuban wild-growing grape liana / Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный

ресурс]. – Краснодар: КубГАУ, 2011. – № 08 (72). С. 272–290. – Режим доступа: <http://ej.kubagro.ru/2011/08/pdf/23.pdf>, 1,188 у.п.л.

16. Трошин Л.П. Современные ампелографические исследования. Морфометрический анализ листовой ампелографической информации // Генетические ресурсы и селекционное обеспечение современного виноградарства. – Новочеркасск: ВНИИВиВ, 2011. – С. 125-132.

17. Головина Н. Е., Трошин Л. П. Системно-когнитивный подход к решению основной задачи ампелографии // [Интерактивная](#) ампелография и селекция винограда. – Краснодар: КубГАУ, 2012. – С. 46-48.

18. Трошин Л.П., Звягин А.С. Технология отбора лучших протоклонов винограда // Технологии производства элитного посадочного материала и виноградной продукции, отбора лучших протоклонов. – Краснодар: АлВи-Дизайн, 2005. – С. 75-95.

19. Трошин Л.П., Животовский Л.А. Методические рекомендации по клоновой селекции винограда на продуктивность / ВНИИВиПП "Магарач". Ин-т общей генетики им. Н.И.Вавилова. - Ялта, 1987. - 36 с.

20. Мандель И.Д. Кластерный анализ. - М.: Финансы и статистика. 1988. – 176с.

21. Леонов В.П. Краткий обзор методов кластерного анализа. Сайт: http://www.biometrica.tomsk.ru/cluster_2.htm
http://www.biometrica.tomsk.ru/cluster_3.htm

22. Леонов В.П. Литература и сайты по кластерному анализу. Сайт: http://www.biometrica.tomsk.ru/cluster_4.htm

23. Сайт Института Космических Исследований РАН: <http://www.iki.rssi.ru/magbase/REFMAN/STATTEXT/modules/stcluan.html#general>

24. Сайт Internet-сообщества закупщиков: http://zakup.vl.ru/132-metodi_klastern.html

25. Баран О.И., Григорьев Ю.А., Жилина Н.М. Алгоритмы и критерии качества кластеризации // Общественное здоровье и здравоохранение: материалы XLV науч.-практ. конф. с международным участием «Гигиена, организация здравоохранения и профпатология» и семинара «Актуальные вопросы современной профпатологии», Новокузнецк, 17-18 ноября 2010 / под ред. В.В.Захаренкова. Кемерово: Примула, 2010. – С. 21-26.

26. Луценко Е.В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе «Эйдос») / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №07(071). С. 528 – 576. – Шифр Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>, 3,062 у.п.л.

27. Луценко Е.В. Количественное измерение сходства-различия клонов винограда по контурам листьев с применением АСК-анализа и системы «Эйдос» / Е.В. Луценко, Л.П. Трошин, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1205 – 1228. – IDA [article ID]: 1161602077. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/77.pdf>, 1,5 у.п.л.

Сканы листьев сортов винограда Виктор, Ливия К, Преображение, Юбилей Новочеркасска:



4. ПРИМЕНЕНИЕ ТЕОРИИ ИНФОРМАЦИИ И КОГНИТИВНЫХ ТЕХНОЛОГИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ГЕНЕТИКИ (на примере вычисления количества информации в генах о признаках и свойствах различных автохтонных сортов винограда)

Общеизвестно, что генетика изучает механизмы изменчивости/наследственности и очень широко пользуется понятием «наследственная информация». При этом генетика под информацией подразумевает содержание генетического кода - структуры молекул ДНК и РНК, входящих в состав клетки живого организма. Генетика изучает механизмы записи, копирования, считывания генетической информации, возможности ее модификации, а также ее влияние на признаки и свойства организма. В разговорном и научном языке прочно закрепились фразы, типа «Гены содержат информацию о признаках/свойствах организма». Парадоксально, но мы не видим попыток определения количества информации, содержащейся в конкретных генах о конкретных фенотипических признаках или свойствах организма. Казалось бы, применение теории информации в генетике является совершенно естественным и напрашивается сама собой. Тем более странно, что практически нет работ, посвященных применению теории информации для решения задач генетики. Данная статья призвана в какой-то степени восполнить этот пробел на примере вычисления количества информации в генах о признаках или свойствах различных сортов винограда. В ней рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и реализующего их программного инструментария – интеллектуальной системы «Эйдос» для решения одной из важных задач генетики: определения количества информации, содержащейся в генах о различных фенотипических признаках/свойствах винограда. Для решения этой задачи выполняются следующие этапы: 1) когнитивно-целевая структуризация предметной области; 2) формализация предметной области, т.е. разработка классификационных и описательных шкал и градаций и обучающей выборки; 3) синтез и верификация информационной модели, отражающей количество информации в генах о фенотипических признаках/свойствах (многопараметрическая типизация); 4) вывод информации о генетической системе детерминации фенотипических признаков/свойств (SWOT-анализ феносвойства); 5) вывод информации о силе и направлении влияния конкретного гена на фенотипические признаки/свойства (SWOT-диаграммы генов); 6) решение задачи системной идентификации фенотипических признаков/свойств по наличию тех или иных генов; 7) количественное определение сходства-различия различных фенотипических признаков/свойств по их системе детерминации генами.

Конкретное фенотипическое свойство (или признак) рассматривается как зашумленный генетический текст, включающий как генетическую информацию об истинном феносвойстве (чистый сигнал), так и шум, искажающий эту информацию, обусловленный случайным воздействием окружающей среды. Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение истинного сигнала.

1. Формулировка проблемы

Общеизвестно, что генетика изучает механизмы наследственности и изменчивости и очень широко пользуется понятием «наследственная информация». При этом генетика под информацией подразумевает содержание генетического кода - структуры молекул ДНК и РНК, входящих в состав клетки живого организма. Генетика изучает механизмы записи, копирования, считывания генетической информации, возможности ее модификации, а также ее влияние на признаки и свойства организма. В разговорном и научном языке прочно закрепились фразы, типа «Гены содержат информацию о признаках и свойствах организма». Парадоксально, но мы не видим систематических попыток определения количества информации, содержащейся в конкретных генах о конкретных фенотипических признаках и свойствах организма. Казалось бы, применение теории информации в генетике является совершенно естественным и напрашивается сама собой. Тем более странно, что практически нет работ, посвященных применению теории информации для решения задач генетики.

В качестве исключения приведем две цитаты из работ 1980 и 1974 годов, в которых совершенно четко осознается роль информации в генетике [46, 47] «Абсолютным критерием живого принято считать наличие определенного количества бит информации, заключенных в информационных молекулах» [46, 47]. «...быстрый рост количества информации дал новое качество – способность к самовоспроизведению (качество, присущее только нуклеиновым кислотам, *обладающим несравненной со всеми другими соединениями нашего мира информационной*

емкостью; качество давшее начало новой форме существования материи)» [46] (курсив наш. авт.). И все же конкретных расчетов количества информации в генах не приводится и в этих работах.

Данная статья призвана представить научной общественности один из возможных вариантов решения этой проблемы на примере вычисления количества информации в генах о признаках и свойствах различных сортов винограда.

2. Идея и предпосылки решения проблемы

По-видимому, сформулированная проблема обусловлена отсутствием доступа у исследователей-генетиков к программному инструментарию, реализующему теорию информации и обеспечивающему решение задач генетики с применением теории информации. Между тем, такой инструментарий существует и находится в полном открытом и бесплатном доступе: это – автоматизированный системно-когнитивный анализ (АСК-анализ) [1], его математическая модель – системная теория информации [2] и реализующий их программный инструментарий – интеллектуальная система «Эйдос» [3].

Этот инструментарий обеспечивает решение одной из важных задач генетики: определения количества информации, содержащейся в генах о различных фенотипических признаках и свойствах винограда. Впервые в четкой и развитой форме эта мысль высказана автором в статье [4].

Для решения этой задачи выполняются следующие этапы АСК-анализа [1]: 1) когнитивно-целевая структуризация предметной области; 2) формализация предметной области, т.е. разработка классификационных и описательных шкал и градаций и обучающей выборки; 3) синтез и верификация информационной модели, отражающей количество информации в генах о фенотипических признаках и свойствах (многопараметрическая типизация); 4) вывод информации о генетической системе детерминации фенотипических признаков

и свойств (SWOT-анализ феносвойства); 5) вывод информации о силе и направлении влияния конкретного гена на фенотипические признаки и свойства (SWOT-диаграммы генов); 6) решение задачи системной идентификации фенотипических признаков/свойств по наличию тех или иных генов; 7) количественное определение сходства-различия различных фенотипических признаков/свойств по их системе детерминации генами. Конкретное фенотипическое свойство (или признак) рассматривается как зашумленный генетический текст, включающий как генетическую информацию об истинном феносвойстве (чистый сигнал), так и шум, искажающий эту информацию, обусловленный случайным воздействием окружающей среды. Программный инструментарий АСК-анализа – интеллектуальная система «Эйдос» обеспечивает подавление шума и выделение истинного сигнала.

Имеется большой опыт успешного применения АСК-анализа для решения широкого круга задач в различных предметных областях.

3. Автоматизированный системно-когнитивный анализ как метод решения проблемы

3.1. Кратко об АСК-анализе

Системный анализ представляет собой современный метод научного познания, общепризнанный метод решения проблем [5, 6]. Однако возможности практического применения системного анализа ограничиваются отсутствием программного инструментария, обеспечивающего его автоматизацию. Существуют разнородные программные системы, автоматизирующие отдельные этапы или функции системного анализа в различных конкретных предметных областях.

Автоматизированный системно-когнитивный анализ (АСК-анализ) представляет собой системный анализ, структурированный по базовым когнитивным операциям (БКО), благодаря чему удалось разработать для него

математическую модель, методику численных расчетов (структуры данных и алгоритмы их обработки), а также реализующую их программную систему – систему «Эйдос» [1-3, 7]. Система «Эйдос» разработана в постановке, не зависящей от предметной области, и имеет ряд программных интерфейсов с внешними данными различных типов [3]. АСК-анализ может быть применен как инструмент, многократно усиливающий возможности естественного интеллекта во всех областях, где используется естественный интеллект. АСК-анализ был успешно применен для решения задач идентификации, прогнозирования, принятия решений и исследования моделируемого объекта путем исследования его модели во многих предметных областях, в частности в экономике, технике, социологии, педагогике, психологии, медицине, экологии, ампелографии, геофизике, энтомологии, криминалистике и др. [8, 9].

3.2. Истоки АСК-анализа

Известно, что системный анализ является одним из общепризнанных в науке методов решения проблем и многими учеными рассматривается вообще как метод научного познания. Однако, как впервые заметил еще в 1984 году проф. И.П. Стабин, на практике применение системного анализа наталкивается на проблему [10]. Суть этой проблемы в том, что обычно системный анализ успешно применяется в сравнительно простых случаях, в которых в принципе можно обойтись и без него, тогда как в действительно сложных ситуациях, когда он действительно чрезвычайно востребован и у него нет альтернатив, сделать это удастся гораздо реже. Проф. И.П. Стабин предложил и путь решения этой проблемы, который он видел в автоматизации системного анализа [10].

Однако путь от идеи до создания программной системы долог и сложен, т.к. включает ряд этапов:

- выбор теоретического математического метода;

– разработка методики численных расчетов, включающей структуры данных в оперативной памяти и внешних баз данных (даталогическую и инфологическую модели) и алгоритмы обработки этих данных;

– разработка программной системы, реализующей эти математические методы и методики численных расчетов.

3.3. Методика АСК-анализа

3.3.1. Предпосылки решения проблемы

Перегудов Ф.И. и Тарасенко Ф.П. в своих основополагающих работах 1989 и 1997 годов [5, 6] подробно рассмотрели математические методы, которые в принципе могли бы быть применены для автоматизации отдельных этапов системного анализа. Однако даже самые лучшие математические методы не могут быть применены на практике без реализующих их программных систем, а путь от математического метода к программной системе долог и сложен. Для этого необходимо разработать численные методы или методики численных расчетов (алгоритмы и структуры данных), реализующие математический метод, а затем разработать программную реализацию системы, основанной на этом численном методе.

В числе первых попыток реальной автоматизации системного анализа следует отметить докторскую диссертацию проф. Симанкова В.С. (2001) [11]. Эта попытка была основана на высокой детализации этапов системного анализа и подборе уже существующих программных систем, автоматизирующих эти этапы. Идея была в том, что чем выше детализация системного анализа, чем мельче этапы, тем проще их автоматизировать. Эта попытка была реализована, однако, лишь для специального случая исследования в области возобновляемой энергетики, т.к. системы оказались различных разработчиков, созданные с помощью различного инструментария и не имеющие программных интерфейсов друг с другом, т.е. не образующие единой автоматизированной системы. Эта попытка, безусловно,

явилась большим шагом по пути, предложенному проф. И.П. Стабиным, но и ее нельзя признать обеспечившей достижение поставленной цели, сформулированной Стабиным И.П. (т.е. создание автоматизированного системного анализа), т.к. она не привела к созданию единой универсальной программной системы, автоматизирующей системный анализ, которую можно было бы применять в различных предметных областях.

Необходимо отметить работы Дж. Клира по системологии и автоматизации решения системных задач, которые внесли большой вклад в автоматизацию системного анализа путем создания и применения универсального решателя системных задач (УРСЗ), реализованного в рамках оригинальной экспертной системы [12, 13]. Однако в экспертной системе применяется продукционная модель знаний, для получения которых от эксперта необходимо участие инженера по знаниям (когнитолога). Этим обусловлены следующие недостатки экспертных систем:

- они генерируют знания каждый раз, когда они необходимы для решения задач, и это может занимать значительно большее время, чем при использовании декларативной формы представления знаний;

- продукционные модели обычно построены на бинарной логике (if then else), что вызывает возможность логического конфликта продукций в процесс логического вывода, что приводит к необратимому останову логического процесса;

- эксперты - люди чаще всего заслуженные и их время и знания стоят очень дорого; поэтому привлечение экспертов для извлечения готовых знаний на длительное время проблематично и обычно эксперт просто физически не может сообщить очень большой объем знаний, а иногда и не хочет этого делать и сообщает неадекватные знания;

- чаще всего эксперты формулируют свои знания неформализуемым путем на основе своей интуиции, опыта и

профессиональной компетенции, т.е. не могут сформулировать свои знания в количественной форме, а пользуются для их формализации порядковыми или даже номинальными шкалами, поэтому экспертные знания являются не очень точными и для их формализации необходим инженер по знаниям (когнитолог).

3.3.2. АСК-анализ как решение проблемы

Автоматизированный системно-когнитивный анализ разработан профессором Е.В. Луценко и предложен в 2002 году [1], хотя разработан он был значительно раньше, причем с программным инструментарием: системой «Эйдос» [1, 3, 7]. Основная идея, позволившая сделать это, состоит в рассмотрении системного анализа как метода познания (отсюда и «когнитивный» от «cognitio» – знание, познание, лат.). Эта идея позволила *структурировать системный анализ не по этапам, как пытались сделать ранее, а по базовым когнитивным операциям системного анализа (БКОСА)*, т.е. таким операциям, к комбинациям которых сводятся остальные. Эти операции образуют минимальную систему, достаточную для описания системного анализа, как метода познания, т.е. конфигуратор. Понятие конфигуратора предложено В.А. Лефевром [14]. В 2002 году Е.В. Луценко был предложен когнитивный конфигуратор [1], включающий 10 базовых когнитивных операций.

Когнитивный конфигуратор:

- 1) присвоение имен;
- 2) восприятие (описание конкретных объектов в форме онтологий, т.е. их признаками и принадлежностью к обобщающим категориям - классам);
- 3) обобщение (синтез, индукция);
- 4) абстрагирование;
- 5) оценка адекватности модели;
- 6) сравнение, идентификация и прогнозирование;
- 7) дедукция и абдукция;
- 8) классификация и генерация конструкторов;

- 9) содержательное сравнение;
- 10) планирование и поддержка принятия управленческих решений.

Каждая из этих операций оказалась достаточно элементарна для формализации и программной реализации.

Компоненты АСК-анализа:

- формализуемая когнитивная концепция и следующий из нее когнитивный конфигуратор;
- теоретические основы, методология, технология и методика АСК-анализа;
- математическая модель АСК-анализа, основанная на системном обобщении теории информации;
- методика численных расчетов, в универсальной форме реализующая математическую модель АСК-анализа, включающая иерархическую структуру данных и 24 детальных алгоритма 10 БКОСА;
- специальное инструментальное программное обеспечение, реализующее математическую модель и численный метод АСК-анализа – Универсальная когнитивная аналитическая система "Эйдос".

Этапы АСК-анализа:

- 1) когнитивно-целевая структуризация предметной области;
- 2) формализация предметной области (конструирование классификационных и описательных шкал и градаций и подготовка обучающей выборки);
- 3) синтез системы моделей предметной области (в настоящее время система «Эйдос» поддерживает 3 статистические модели и 7 системно-когнитивных моделей (моделей знаний);
- 4) верификация (оценка достоверности) системы моделей предметной области;
- 5) повышение качества системы моделей;

6) решение задач идентификации, прогнозирования и поддержки принятия решений;

7) исследование моделируемого объекта путем исследования его моделей является корректным, если модель верно отражает моделируемый объект и включает: кластерно-конструктивный анализ классов и факторов; содержательное сравнение классов и факторов; изучение системы детерминации состояний моделируемого объекта; нелокальные нейроны и интерпретируемые нейронные сети прямого счета; классические когнитивные модели (когнитивные карты); интегральные когнитивные модели (интегральные когнитивные карты), прямые обратные SWOT-диаграммы; когнитивные функции и т.д.

Математические аспекты АСК-анализа

Математическая модель АСК-анализ основана на теории информации, точнее на системной теории информации (СТИ), предложенной Е.В. Луценко [1, 2, 3]¹⁰. Это значит, что *в АСК-анализе все факторы рассматриваются с одной единственной точки зрения: сколько информации содержится в их значениях о переходе объекта, на который они действуют, в определенное состояние, и при этом сила и направление влияния всех значений факторов на объект измеряется в одних общих для всех факторов единицах измерения: единицах количества информации [8, 9].*

Это напоминает подход Дугласа Хаббарда [15], но, в отличие от него, имеет открытый универсальный программный инструментарий (систему «Эйдос»), разработанный в постановке, не зависящей от предметной области [1-3]. К тому же на систему «Эйдос» уже в 1994 году было три патента РФ [3, 16¹¹], а первые акты ее внедрения датируются 1987 годом [1, 3]¹², тогда как основная работа Дугласа Хаббарда [15] появилась лишь в 2009

¹⁰ Математическая модель АСК-анализа описана в ряде работ:
http://elibrary.ru/author_items.asp?authorid=123162

¹¹ См., например: <http://lc.kubagro.ru/aidos/index.htm>

¹² <http://lc.kubagro.ru/aidos/aidos02/PR-4.htm>

году. Это означает, что идеи АСК-анализа не только появились, но и были доведены до программной реализации в универсальной форме и применены в различных предметных областях на 22 с лишним года *раньше* появления работ Дугласа Хаббарда.

Поэтому АСК-анализ обеспечивает корректную сопоставимую обработку числовых и нечисловых данных, представленных в разных типах измерительных шкал и разных единицах измерения [8, 9]. Метод АСК-анализа является устойчивым непараметрическим методом, обеспечивающим создание моделей больших размерностей при неполных и зашумленных исходных данных о сложном нелинейном динамичном объекте управления. Этот метод является чуть ли не единственным на данный момент, обеспечивающим многопараметрическую типизацию и системную идентификацию методов, инструментарий которого (интеллектуальная система «Эйдос») находится в полном открытом бесплатном доступе [3, 16]¹³.

3.4. Некоторые результаты применения АСК-анализа в различных предметных областях

Метод системно-когнитивного анализа и его программный инструментарий интеллектуальная система "Эйдос" были успешно применены при проведении 6 докторских и 7 кандидатских диссертационных работ в ряде различных предметных областей по экономическим, техническим, психологическим и медицинским наукам.

АСК-анализ был успешно применены при выполнении десятков грантов РФФИ и РГНФ различной направленности за длительный период - с 2002 года по настоящее время (2016 год).

По проблематике АСК-анализа издана 21 монография, получено 29 патентов на системы искусственного интеллекта, их подсистемы, режимы и приложения, опубликовано более 200 статей в изданиях, входящих в Перечень ВАК РФ (по данным

¹³ <http://lc.kubagro.ru/aidos/Aidos-X.htm>

РИНЦ). В одном только Научном журнале КубГАУ (входит в Перечень ВАК РФ с 26-го марта 2010 года) автором АСК-анализа Луценко Е.В. опубликовано 186 статей общим объёмом 321,559 у.п.л., в среднем 1,729 у.п.л. на одну статью.

По этим публикациям, грантам и диссертационным работам видно, что АСК-анализ уже был успешно применен в следующих предметных областях и научных направлениях: экономика (региональная, отраслевая, предприятий, прогнозирование фондовых рынков), социология, эконометрика, биометрия, педагогика (создание педагогических измерительных инструментов и их применение), психология (личности, экстремальных ситуаций, профессиональных и учебных достижений, разработка и применение профессиограмм), сельское хозяйство (прогнозирование результатов применения агротехнологий, принятие решений по выбору рациональных агротехнологий и микрон зон выращивания), экология, ампелография, геофизика (глобальное и локальное прогнозирование землетрясений, параметров магнитного поля Земли, движения полюсов Земли), климатология (прогнозирование Эль-Ниньо и Ла-Нинья), возобновляемая энергетика, мелиорация и управление мелиоративными системами, криминалистика, энтомология и ряд других областей.

АСК-анализ вызывает большой интерес во всем мире. Сайт автора АСК-анализа [16] посетило около 500 тыс. посетителей с уникальными IP-адресами со всего мира. Еще около 500 тыс. посетителей открывали статьи по АСК-анализу в Научном журнале КубГАУ.

Необходимо отметить, что в развитии различных теоретических основ и практических аспектов АСК-анализа приняли участие многие ученые: д.э.н., к.т.н., проф. Луценко Е.В., Засл. деятель науки РФ, д.т.н., проф. Лойко В.И., к.ф.-м.н., Ph.D., проф., Трунев А.П. (Канада), д.э.н., д.т.н., к.ф.-м.н., проф. Орлов А.И., к.т.н., доц. Коржаков В.Е., д.э.н., проф. Барановская

Т.П., д.э.н., к.т.н., проф. Ермоленко В.В., к.пс.н. Наприев И.Л., к.пс.н., доц. Некрасов С.Д., к.т.н., доц. Лаптев В.Н., к.пс.н, доц. Третьяк В.Г., к.пс.н., Щукин Т.Н., д.т.н., проф. Симанков В.С., д.э.н., проф. Ткачев А.Н., д.т.н., проф. Сафронова Т.И., д.э.н., доц. Горпинченко К.Н., к.э.н., доц. Макаревич О.А., к.э.н., доц. Макаревич Л.О., к.м.н. Сергеева Е.В. (Фомина Е.В.), Бандык Д.К. (Белоруссия), Чередниченко Н.А., к.ф.-м.н. Артемов А.А., д.э.н., проф. Крохмаль В.В., д.т.н., проф. Рябцев В.Г., к.т.н., доц. Марченко А.Ю., д.т.н., проф. Фролов В.Ю., д.ю.н, проф. Швец С.В., Засл. деятель науки Кубани, д.б.н., проф. Трошин Л.П., Засл. изобр. РФ, д.т.н., проф. Серга Г.В., Сергеев А.С., д.б.н., проф. Стрельников В.В. и другие.

4. Решение проблемы

В данном разделе рассмотрим решение поставленной проблемы путем последовательного выполнения различных этапов АСК-анализа [1].

4.1. Когнитивно-целевая структуризация предметной области

На этом этапе экспертом неформализуемым путем на основе опыта, интуиции и профессиональной компетенции определяется, что является причинами (факторами), а что следствиями (результатами действия факторов) в моделируемой предметной области.

В нашем случае вполне очевидно, что как факторы целесообразно рассматривать количество резервных генов различных видов, а как результаты их действия – различные фенотипические признаки или свойства различных сортов винограда.

В таблице 1 представлены классификационные шкалы, с помощью которых будут кодироваться фенотипические признаки, а в таблице 2 – описательные шкалы, т.е. резервные гены на 1-й и 2-й спиралях ДНК:

Таблица 1. – Классификационные шкалы

| Код | Наименование |
|-----|---|
| 1 | ТИП ЦВЕТКА (151) |
| 2 | СРЕДНЯЯ МАССА ГРОЗДИ, Г |
| 3 | УРОЖАЙНОСТЬ, Ц/ГА |
| 4 | САХАРИСТОСТЬ, Г/100 КУБ. СМ |
| 5 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л |
| 6 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072) |
| 7 | ПУЗЫРЧАТОСТЬ ЛИСТЬЕВ (075) |
| 8 | ВЕРХНИЕ ВЫРЕЗКИ ЛИСТЬЕВ (094-1) |
| 9 | НИЖНИЕ ВЫРЕЗКИ ЛИСТЬЕВ (094-2) |
| 10 | ОПУШЕНИЕ ЛИСТЬЕВ (084) |
| 11 | РАЗМЕР ГРОЗДИ (202) |
| 12 | ПЛОТНОСТЬ ГРОЗДИ (204) |
| 13 | РАЗМЕР ЯГОДЫ (220) |
| 14 | ФОРМА ЯГОДЫ (223) |
| 15 | ОКРАСКА ЯГОДЫ (225) |
| 16 | АРОМАТ ЯГОДЫ (236) |
| 17 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603) |
| 18 | ПЕРИОД СОЗРЕВАНИЯ (629) |
| 19 | ПРОДОЛЖИТЕЛЬНОСТЬ ПРОДУКЦИОННОГО ПЕРИОДА, ДНИ |
| 20 | РОСТ ПОБЕГОВ (351) |

Таблица 2. – Описательные шкалы

| Код | Наименование | Код | Наименование | Код | Наименование |
|-----|--------------|-----|--------------|-----|--------------|
| 1 | VRZAG47-2 | 18 | VVMD25-1 | 35 | VVIV37-2 |
| 2 | VRZAG47-1 | 19 | VVIP60-2 | 36 | VVIV37-1 |
| 3 | VVS2-2 | 20 | VVIP60-1 | 37 | VVIN54-2 |
| 4 | VVS2-1 | 21 | VVIB01-2 | 38 | VVIN54-1 |
| 5 | VVMD7-2 | 22 | VVIB01-1 | 39 | VVMD24-2 |
| 6 | VVMD7-1 | 23 | VRZAG83-2 | 40 | VVMD24-1 |
| 7 | VVMD5-2 | 24 | VRZAG83-1 | 41 | VVIV67-2 |
| 8 | VVMD5-1 | 25 | VVMD27-2 | 42 | VVIV67-1 |
| 9 | VRZAG62-2 | 26 | VVMD27-1 | 43 | VVIN73-2 |
| 10 | VRZAG62-1 | 27 | VVMD21-2 | 44 | VVIN73-1 |
| 11 | VRZAG79-2 | 28 | VVMD21-1 | 45 | VMC4F3.1-2 |
| 12 | VRZAG79-1 | 29 | VMC1B11-2 | 46 | VMC4F3.1-1 |
| 13 | VVMD28-2 | 30 | VMC1B11-1 | 47 | VVIN16-2 |
| 14 | VVMD28-1 | 31 | VVIQ52-2 | 48 | VVIN16-1 |
| 15 | VVMD32-2 | 32 | VVIQ52-1 | 49 | VVIP31-2 |
| 16 | VVMD32-1 | 33 | VRZAG67-2 | 50 | VVIP31-1 |
| 17 | VVMD25-2 | 34 | VRZAG67-1 | | |

Последующие этапы АСК-анализа автоматизированы в его инструментарии – интеллектуальной системе «Эйдос».

4.2. Формализация предметной области, т.е. разработка классификационных и описательных шкал и градаций и обучающей выборки

На этапе формализации предметной области разрабатываются классификационные и описательные шкалы и градации и с их помощью кодируются исходные данные, в результате чего формируется обучающая выборка.

В соответствии с результатами когнитивно-целевой структуризации предметной области исходные данные представлены в виде Excel-таблицы (таблица 3). Полностью в статье таблица 3 не приводится из-за ее размерности: в ней 90 строк и 71 колонка. Но она (вместе с другими материалами по статье) приводится по ссылке: <http://ej.kubagro.ru/2016/07/upload/03.zip>.

Структура таблицы 3 соответствует требованиям системы «Эйдос», приведенным на рисунке 1, точнее требованиям ее универсального программного интерфейса импорта данных из внешних баз данных (режим 2.3.2.2).

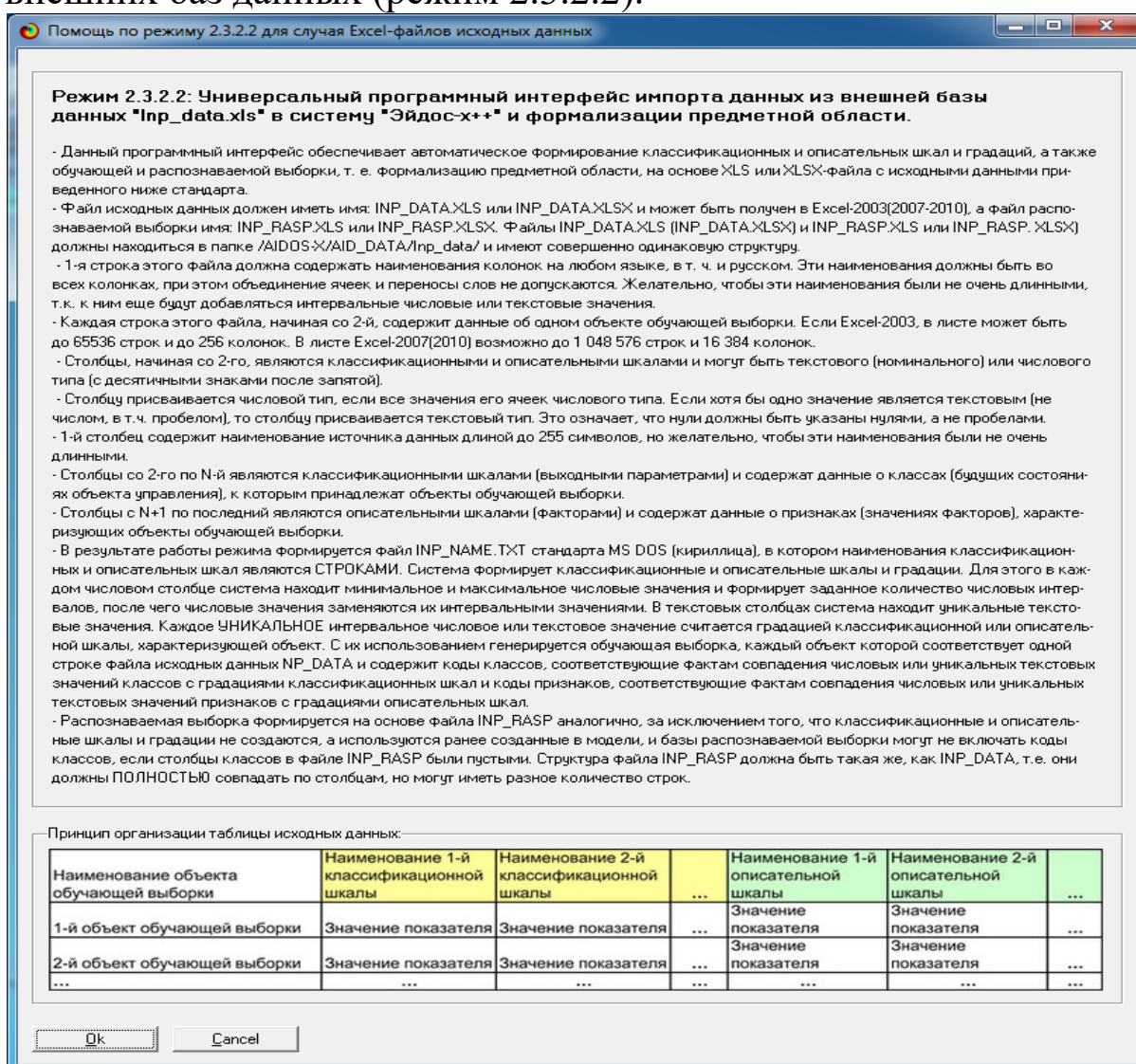


Рисунок 1. Требования к внешнему файлу исходных данных

Таблица 3. – Исходные данные (фрагмент)

| Наименование сорта | Тип цветка (151) | Средняя масса грозди, г | Урожайность, ц/га | Сахаристость, г/100 куб. см | Кислотность титруемая, г/л | Расеченность листьев (072) | Пузырчатость листьев (075) | Вырезки верхние листьев (094-1) | Вырезки нижние листьев (094-2) | Опушение листьев (084) | Грозди размер (202) | Плотность грозди (204) | Ягоды размер (220) | Ягоды форма (223) | Ягоды окраска (225) | Ягоды аромат (236) | Направление использования (603) | Период созре-вания (629) | Продолжительность продукционного периода, дни | Рост побегов (351) | VrZAG47-2 | VrZAG47-1 | VVS2-2 | VVS2-1 | VVMD7-2 | VVMD7-1 | VVMD5-2 |
|---------------------|------------------|-------------------------|-------------------|-----------------------------|----------------------------|----------------------------|----------------------------|---------------------------------|--------------------------------|------------------------|---------------------|------------------------|--------------------|-------------------|---------------------|--------------------|---------------------------------|--------------------------|---|--------------------|-----------|-----------|--------|--------|---------|---------|---------|
| Аджем мискет | :4 | 40,0 | 25,0 | 21,5 | 3,0 | :7 | :3 | :7 | :1 | :3 | :3 | :3 | :5 | :4 | :1 | :2 | :1 | :6 | 157 | :5 | 182 | 188 | 143 | 145 | 239 | 239 | 242 |
| Аг изюм | :3 | 220,0 | 120,0 | 18,5 | 8,0 | :7 | :5 | :7 | :7 | :6 | :5 | :3 | :7 | :3 | :1 | :4 | :1 | :5 | 130 | :5 | 180 | 195 | 135 | 145 | 249 | 249 | 230 |
| Аккерманский черный | :3 | 95,0 | 55,0 | 20,5 | 6,5 | :3 | :5 | :3 | :3 | :4 | :3 | :7 | :5 | :2 | :7 | :1 | :4 | :6 | 152 | :7 | 182 | 195 | 133 | 143 | 239 | 249 | 240 |
| Аклык желтый | :3 | 180,0 | 35,0 | 17,5 | 6,5 | :3 | :5 | :3 | :1 | :3 | :5 | :5 | :7 | :4 | :1 | :1 | :3 | :5 | 142 | :5 | 195 | 195 | 143 | 145 | 239 | 249 | 236 |
| Аленький | :4 | 60,0 | 30,0 | 21,9 | 10,7 | :5 | :5 | :5 | :5 | :3 | :5 | :3 | :3 | :7 | :1 | :4 | :5 | 143 | :5 | 180 | 195 | 133 | 135 | 239 | 249 | 228 | |
| Алый терский | :3 | 225,0 | 175,0 | 18,0 | 6,7 | :7 | :9 | :7 | :7 | :5 | :5 | :3 | :5 | :3 | :6 | :1 | :4 | :7 | 146 | :7 | 182 | 195 | 135 | 143 | 245 | 249 | 238 |
| Амет аджи ибрам | :3 | 110,0 | 65,0 | 17,5 | 5,5 | :3 | :5 | :3 | :1 | :8 | :5 | :5 | :5 | :3 | :1 | :1 | :4 | :6 | 152 | :5 | 180 | 182 | 143 | 143 | 239 | 239 | 236 |
| Асма | :3 | 400,0 | 95,0 | 17,1 | 7,6 | :1 | :5 | :5 | :3 | :1 | :7 | :7 | :7 | :7 | :1 | :1 | :1 | :8 | 160 | :7 | 186 | 195 | 149 | 151 | 243 | 247 | 234 |
| Баят капы | :3 | 312,0 | 148,0 | 17,9 | 10,0 | :5 | :5 | :5 | :5 | :9 | :7 | :9 | :5 | :3 | :6 | :8 | :4 | :5 | 141 | :7 | 182 | 182 | 135 | 145 | 239 | 243 | 236 |
| Богос зерва | :3 | 135,0 | 60,0 | 19,0 | 5,5 | :5 | :4 | :7 | :5 | :8 | :5 | :7 | :3 | :3 | :1 | :1 | :4 | :5 | 151 | :5 | 182 | 182 | 145 | 149 | 239 | 239 | 240 |
| Бор кара | :3 | 286,0 | 150,0 | 16,0 | 5,5 | :7 | :5 | :7 | :7 | :7 | :7 | :5 | :5 | :3 | :6 | :1 | :4 | :7 | 144 | :7 | 176 | 186 | 141 | 143 | 239 | 253 | 238 |
| Борю изюм | :3 | 225,0 | 170,0 | 14,5 | 4,5 | :5 | :1 | :5 | :5 | :3 | :5 | :5 | :5 | :3 | :1 | :1 | :1 | :5 | 143 | :5 | 190 | 195 | 135 | 137 | 239 | 249 | 236 |
| Брусковатенький | :4 | 100,0 | 17,5 | 22,1 | 8,5 | :5 | :5 | :5 | :3 | :1 | :3 | :3 | :3 | :7 | :1 | :4 | :5 | 140 | :5 | 180 | 182 | 129 | 143 | 239 | 243 | 230 | |
| Будай шули | :3 | 325,0 | 275,0 | 17,5 | 7,4 | :7 | :5 | :7 | :7 | :7 | :7 | :5 | :7 | :7 | :1 | :4 | :3 | :7 | 147 | :6 | 180 | 186 | 141 | 143 | 239 | 247 | 238 |
| Буланый белый | :3 | 125,0 | 80,0 | 20,2 | 7,2 | :5 | :5 | :5 | :3 | :1 | :5 | :5 | :3 | :2 | :1 | :1 | :3 | :5 | 140 | :7 | 180 | 184 | 133 | 145 | 239 | 239 | 240 |
| Чауш чабан | :3 | 250,0 | 100,0 | 17,0 | 5,0 | :7 | :5 | :7 | :7 | :2 | :6 | :5 | :6 | :3 | :1 | :4 | :1 | :6 | 150 | :7 | 180 | 190 | 135 | 143 | 243 | 247 | 230 |
| Чол бер | :4 | 277,0 | 151,0 | 15,0 | 6,9 | :3 | :4 | :3 | :3 | :1 | :7 | :5 | :7 | :3 | :6 | :1 | :3 | :7 | 150 | :5 | 180 | 182 | 133 | 135 | 239 | 247 | 238 |
| Чоркуца розовэ | :3 | 240,0 | 125,0 | 17,0 | 7,0 | :7 | :4 | :7 | :7 | :8 | :5 | :5 | :3 | :3 | :2 | :4 | :4 | :6 | 152 | :7 | 182 | 195 | 133 | 135 | 239 | 239 | 230 |
| Демир кара | :3 | 165,0 | 65,0 | 20,5 | 7,2 | :7 | :4 | :7 | :3 | :3 | :5 | :7 | :5 | :2 | :7 | :7 | :4 | :5 | 151 | :5 | 180 | 188 | 135 | 145 | 239 | 253 | 234 |
| Докур | :3 | 304,0 | 70,0 | 16,0 | 5,5 | :7 | :1 | :7 | :7 | :1 | :7 | :5 | :6 | :4 | :1 | :1 | :1 | :7 | 150 | :5 | 186 | 186 | 135 | 151 | 249 | 249 | 230 |
| Дубут | :3 | 325,0 | 150,0 | 17,4 | 6,0 | :3 | :4 | :5 | :3 | :8 | :7 | :5 | :7 | :3 | :1 | :4 | :1 | :7 | 143 | :7 | 182 | 195 | 135 | 145 | 239 | 259 | 230 |
| Джагар | :3 | 308,0 | 249,0 | 17,0 | 6,2 | :7 | :4 | :7 | :7 | :3 | :7 | :5 | :7 | :3 | :2 | :1 | :3 | :7 | 148 | :7 | 182 | 186 | 135 | 143 | 245 | 259 | 236 |
| Джеват кара | :3 | 172,0 | 120,0 | 19,2 | 4,2 | :5 | :7 | :5 | :3 | :8 | :5 | :5 | :5 | :3 | :6 | :1 | :4 | :6 | 153 | :5 | 180 | 184 | 143 | 145 | 239 | 249 | 230 |

В 1-й колонке таблицы приведены наименования сортов винограда. В колонках со 2-й по 21-ю приведены классификационные шкалы, с помощью градаций которых описывается степень выраженности фенотипических признаков и свойств. При этом и в наименованиях самих шкал, и в наименованиях их градаций используется общепринятая международная система кодирования фенотипических признаков и свойств (выделены светло-голубым фоном) [17]. В колонках с 22-й по 71-ю приведены наименования микросателлитных генов на 1-й и 2-й спиральях ДНК. В таблице 1 допускаются колонки с текстовым и числовым типом данных [26-45].

Резервные гены выбраны авторами для исследования по двум причинам:

- во-первых, потому, что по ним была информация;
- во-вторых, потому, что традиционно считается, что они не связаны с фенотипическими признаками и свойствами, что на самом деле не так, как показало данное исследование. Для ввода исходных из таблицы 3 в систему «Эйдос» и формирования классификационных и описательных шкал и градаций и обучающей выборки запускаем режим 2.3.2.2 с параметрами, приведенными на рисунке 2.

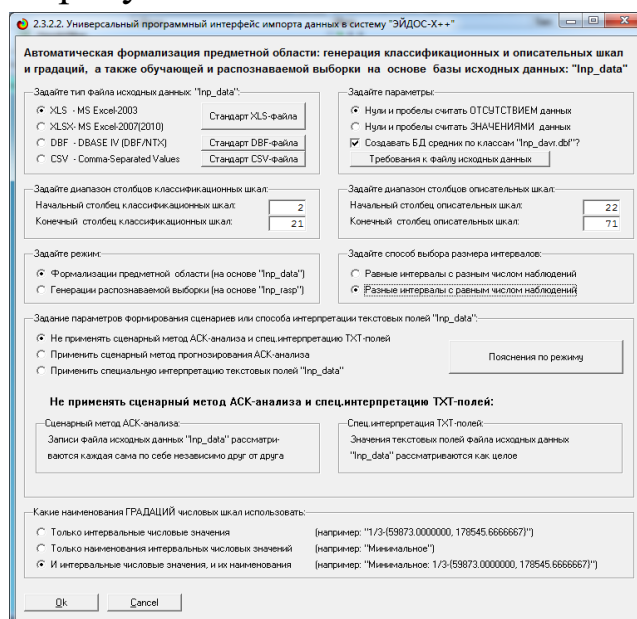


Рисунок 2. Параметры программного интерфейса импорта данных

Во внутреннем калькуляторе (рисунок 3) задаем по 5 градаций в числовых классификационных и описательных шкалах. В результате работы программного интерфейса получаем классификационные шкалы и градации (таблица 4), описательные шкалы и градации (таблица 5) и обучающую выборку или эвентологическую базу данных (таблица 6). Обучающая выборка (таблица 6) представляет собой исходные данные (таблица 1), закодированные с помощью классификационных и описательных шкал и градаций (таблицы 4, 5).

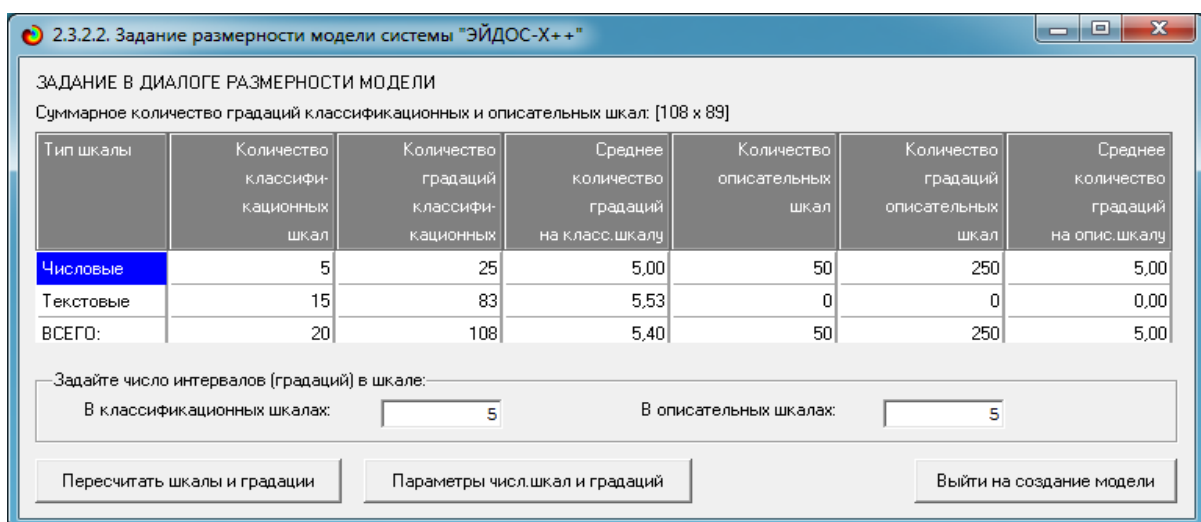


Рисунок 3. Внутренний калькулятор программного интерфейса импорта данных из внешних баз данных

В течение 3-х секунд формируются классификационные и описательные шкалы и градации и обучающая выборка (рисунок 4):

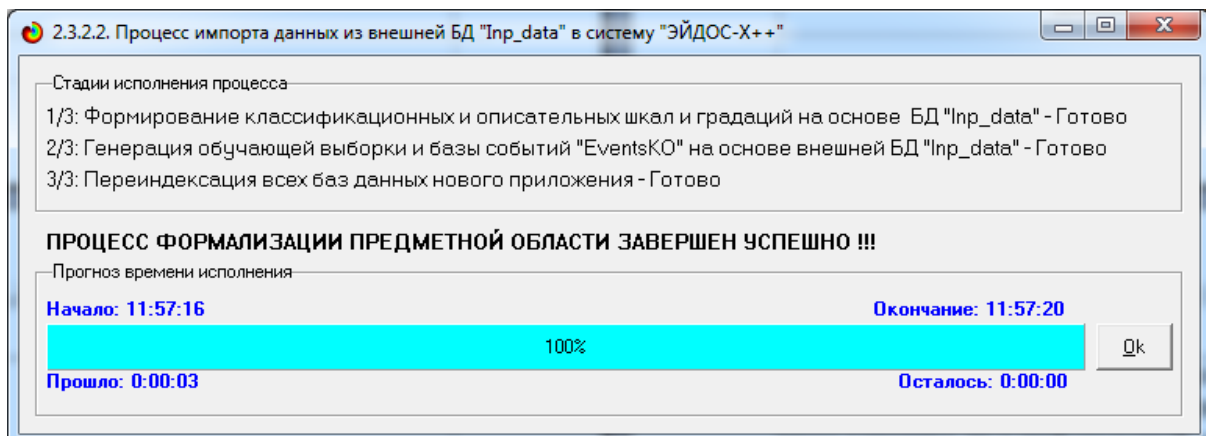


Рисунок 4. Отображение процесса формализации предметной области

В таблицах 4, 5, 6 и последующих часто приводятся лишь фрагменты таблиц из-за их большой размерности ограниченности объема статьи. Но полностью модель вместе с системой можно скачать по ссылке: <http://ej.kubagro.ru/2016/07/upload/03.rar>.

Таблица 4 – Классификационные шкалы и градации (фрагмент)

| Код | Наименование класса |
|-----|---|
| 1 | ТИП ЦВЕТКА (151)-1/2-:3 |
| 2 | ТИП ЦВЕТКА (151)-2/2-:4 |
| 3 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Очень малое: 1/5-{27.0000000, 112.0000000} |
| 4 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Малое: 2/5-{112.0000000, 173.0000000} |
| 5 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Среднее: 3/5-{173.0000000, 220.0000000} |
| 6 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Большое: 4/5-{220.0000000, 280.0000000} |
| 7 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Очень большое: 5/5-{280.0000000, 400.0000000} |
| 8 | УРОЖАЙНОСТЬ, Ц/ГА-Очень малое: 1/5-{9.0000000, 60.0000000} |
| 9 | УРОЖАЙНОСТЬ, Ц/ГА-Малое: 2/5-{60.0000000, 95.0000000} |
| 10 | УРОЖАЙНОСТЬ, Ц/ГА-Среднее: 3/5-{95.0000000, 120.0000000} |
| 11 | УРОЖАЙНОСТЬ, Ц/ГА-Большое: 4/5-{120.0000000, 170.0000000} |
| 12 | УРОЖАЙНОСТЬ, Ц/ГА-Очень большое: 5/5-{170.0000000, 409.0000000} |
| 13 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Очень малое: 1/5-{14.5000000, 17.0000000} |
| 14 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Малое: 2/5-{17.0000000, 18.0000000} |
| 15 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Среднее: 3/5-{18.0000000, 18.7000000} |
| 16 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Большое: 4/5-{18.7000000, 20.5000000} |
| 17 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Очень большое: 5/5-{20.5000000, 24.0000000} |
| 18 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л-Очень малое: 1/5-{2.8000000, 5.5000000} |
| 19 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л-Малое: 2/5-{5.5000000, 6.4000000} |
| 20 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л-Среднее: 3/5-{6.4000000, 6.9000000} |
| 21 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л-Большое: 4/5-{6.9000000, 7.8000000} |
| 22 | КИСЛОТНОСТЬ ТИТРУЕМАЯ, Г/Л-Очень большое: 5/5-{7.8000000, 10.7000000} |
| 23 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-1/6-:1 |
| 24 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-2/6-:3 |
| 25 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-3/6-:5 |
| 26 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-4/6-:6 |
| 27 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-5/6-:7 |
| 28 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-6/6-:9 |
| 29 | ПУЗЫРЧАТОСТЬ ЛИСТЬЕВ (075)-1/6-:1 |
| 30 | ПУЗЫРЧАТОСТЬ ЛИСТЬЕВ (075)-2/6-:3 |

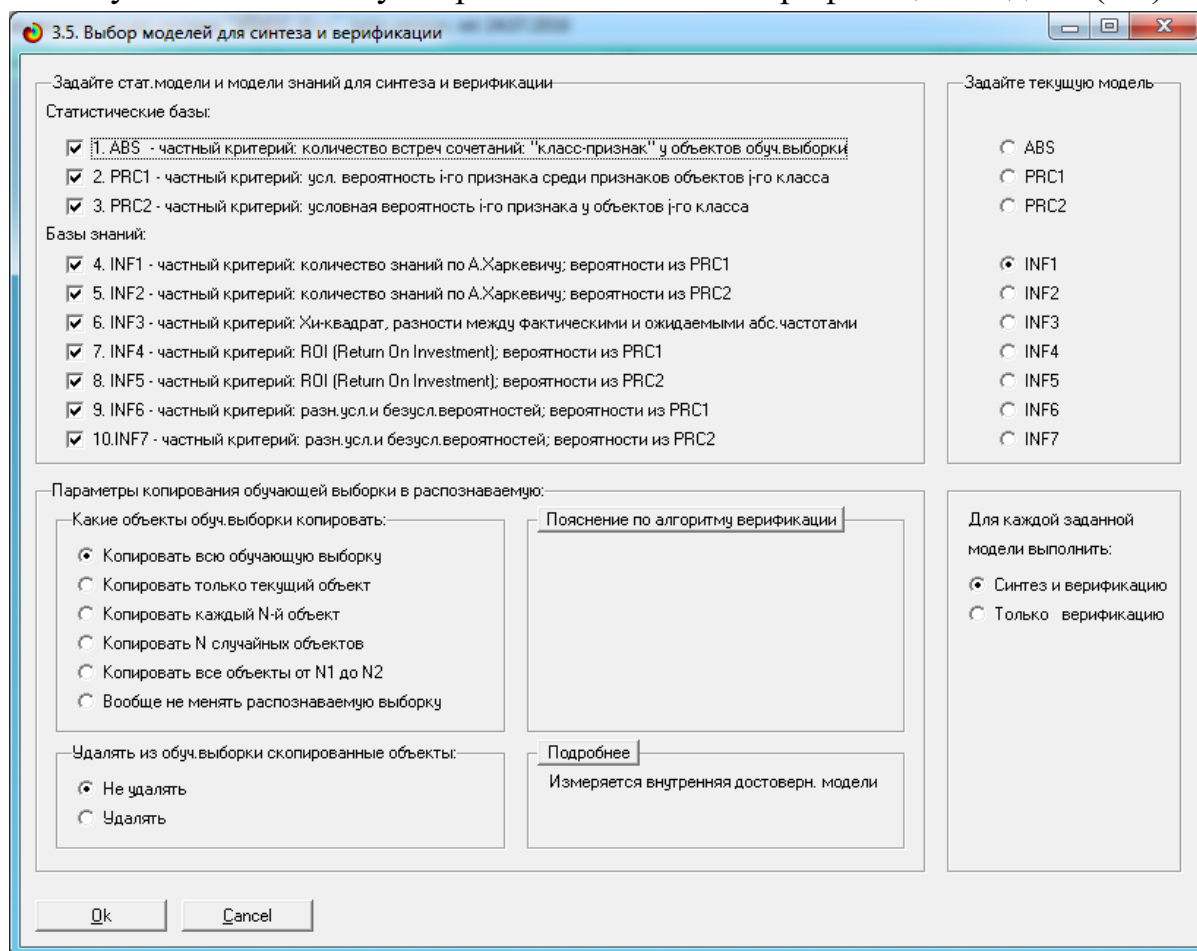
Таблица 5. – Описательные шкалы и градации (фрагмент)

| Код | Наименование |
|-----|---|
| 1 | VRZAG47-2-Очень малое: 1/5-{176.0000000, 180.0000000} |
| 2 | VRZAG47-2-Малое: 2/5-{180.0000000, 182.0000000} |
| 3 | VRZAG47-2-Среднее: 3/5-{182.0000000, 182.0000000} |
| 4 | VRZAG47-2-Большое: 4/5-{182.0000000, 186.0000000} |
| 5 | VRZAG47-2-Очень большое: 5/5-{186.0000000, 195.0000000} |
| 6 | VRZAG47-1-Очень малое: 1/5-{180.0000000, 184.0000000} |
| 7 | VRZAG47-1-Малое: 2/5-{184.0000000, 186.0000000} |
| 8 | VRZAG47-1-Среднее: 3/5-{186.0000000, 190.0000000} |
| 9 | VRZAG47-1-Большое: 4/5-{190.0000000, 195.0000000} |
| 10 | VRZAG47-1-Очень большое: 5/5-{195.0000000, 195.0000000} |
| 11 | VVS2-2-Очень малое: 1/5-{123.0000000, 133.0000000} |
| 12 | VVS2-2-Малое: 2/5-{133.0000000, 135.0000000} |
| 13 | VVS2-2-Среднее: 3/5-{135.0000000, 135.0000000} |
| 14 | VVS2-2-Большое: 4/5-{135.0000000, 143.0000000} |
| 15 | VVS2-2-Очень большое: 5/5-{143.0000000, 149.0000000} |
| 16 | VVS2-1-Очень малое: 1/5-{135.0000000, 141.0000000} |
| 17 | VVS2-1-Малое: 2/5-{141.0000000, 143.0000000} |
| 18 | VVS2-1-Среднее: 3/5-{143.0000000, 143.0000000} |
| 19 | VVS2-1-Большое: 4/5-{143.0000000, 145.0000000} |
| 20 | VVS2-1-Очень большое: 5/5-{145.0000000, 155.0000000} |
| 21 | VVMD7-2-Очень малое: 1/5-{231.0000000, 239.0000000} |
| 22 | VVMD7-2-Малое: 2/5-{239.0000000, 239.0000000} |
| 23 | VVMD7-2-Среднее: 3/5-{239.0000000, 239.0000000} |
| 24 | VVMD7-2-Большое: 4/5-{239.0000000, 247.0000000} |
| 27 | VVMD7-1-Малое: 2/5-{243.0000000, 247.0000000} |
| 28 | VVMD7-1-Среднее: 3/5-{247.0000000, 249.0000000} |
| 29 | VVMD7-1-Большое: 4/5-{249.0000000, 249.0000000} |
| 30 | VVMD7-1-Очень большое: 5/5-{249.0000000, 265.0000000} |

4.3. Синтез и верификация информационной модели, отражающей количество информации в генах о фенотипических свойствах (многопараметрическая типизация)

После формализации предметной области запускаем режим 3.5, обеспечивающий синтез и верификацию моделей с параметрами, приведенными на рисунке 5.

Рисунок 5. Окно запуска режима синтеза и верификации модели (3.5)



Частные модели ABS, PRC#, INF#, отличаются друг друга частными критериями знаний [8, 9] (таблица 7).

Таблица 7. – Частные критерии знаний, используемые в настоящее время в АСК-анализе и системе «Эйдос-Х++»

| Наименование модели знаний и частный критерий | Выражение для частного критерия | |
|---|--|--|
| | через относительные частоты | через абсолютные частоты |
| INF1 , частный критерий: количество знаний по А.Харкевичу, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу. Относительная частота того, что если у объекта j -го класса обнаружен признак, то это i -й признак | $I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$ | $I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_iN_j}$ |
| INF2 , частный критерий: количество знаний по А.Харкевичу, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу. Относительная частота того, что если предъявлен объект j -го класса, то у него будет обнаружен i -й признак. | $I_{ij} = \Psi \times \text{Log}_2 \frac{P_{ij}}{P_i}$ | $I_{ij} = \Psi \times \text{Log}_2 \frac{N_{ij}N}{N_iN_j}$ |
| INF3 , частный критерий: Хи-квадрат: разности между фактическими и теоретически ожидаемыми абсолютными частотами | --- | $I_{ij} = N_{ij} - \frac{N_iN_j}{N}$ |
| INF4 , частный критерий: ROI - Return On Investment, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу ¹⁴ | $I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$ | $I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$ |
| INF5 , частный критерий: ROI - Return On Investment, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу | $I_{ij} = \frac{P_{ij}}{P_i} - 1 = \frac{P_{ij} - P_i}{P_i}$ | $I_{ij} = \frac{N_{ij}N}{N_iN_j} - 1$ |
| INF6 , частный критерий: разность условной и безусловной относительных частот, 1-й вариант расчета относительных частот: N_j – суммарное количество признаков по j -му классу | $I_{ij} = P_{ij} - P_i$ | $I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$ |
| INF7 , частный критерий: разность условной и безусловной относительных частот, 2-й вариант расчета относительных частот: N_j – суммарное количество объектов по j -му классу | $I_{ij} = P_{ij} - P_i$ | $I_{ij} = \frac{N_{ij}}{N_j} - \frac{N_i}{N}$ |

Обозначения:

i – значение прошлого параметра;

j – значение будущего параметра;

N_{ij} – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра;

¹⁴ Применение предложено Л.О. Макаревич

M – суммарное число значений всех прошлых параметров;
 W – суммарное число значений всех будущих параметров;
 N_i – количество встреч i -м значения прошлого параметра по всей выборке;
 N_j – количество встреч j -го значения будущего параметра по всей выборке;
 N – количество встреч j -го значения будущего параметра при i -м значении прошлого параметра по всей выборке;
 I_{ij} – частный критерий знаний: количество знаний в факте наблюдения i -го значения прошлого параметра о том, что объект перейдет в состояние, соответствующее j -му значению будущего параметра;
 Ψ – нормировочный коэффициент (Е.В. Луценко, 2002), преобразующий количество информации в формуле А.Харкевича в биты и обеспечивающий для нее соблюдение принципа соответствия с формулой Р.Хартли;
 P_i – безусловная относительная частота встречи i -го значения прошлого параметра в обучающей выборке;
 P_{ij} – условная относительная частота встречи i -го значения прошлого параметра при j -м значении будущего параметра.

Все эти способы метризации с применением 7 частных критериев знаний (таблица 7) реализованы в системно-когнитивном анализе и интеллектуальной системе «Эйдос» и обеспечивают сопоставление градациям всех видов шкал числовых значений, имеющих смысл количества информации в градации о принадлежности объекта к классу. Поэтому является корректным применение интегральных критериев, включающих операции умножения и суммирования, для обработки числовых значений, соответствующих градациям шкал. Это позволяет единообразно и сопоставимо обрабатывать эмпирические данные, полученные с помощью любых типов шкал, применяя при этом все математические операции.

На рисунке 6 приведено окно с отображением процесса исполнения режима 3.5. Из этого рисунка видно, что синтез и верификация модели была выполнена за 6 минут 48 секунд (процессор i7). Отметим, что встречаются задачи интеллектуальной обработки больших данных, в которых этот режим может выполняться неделю и даже более.

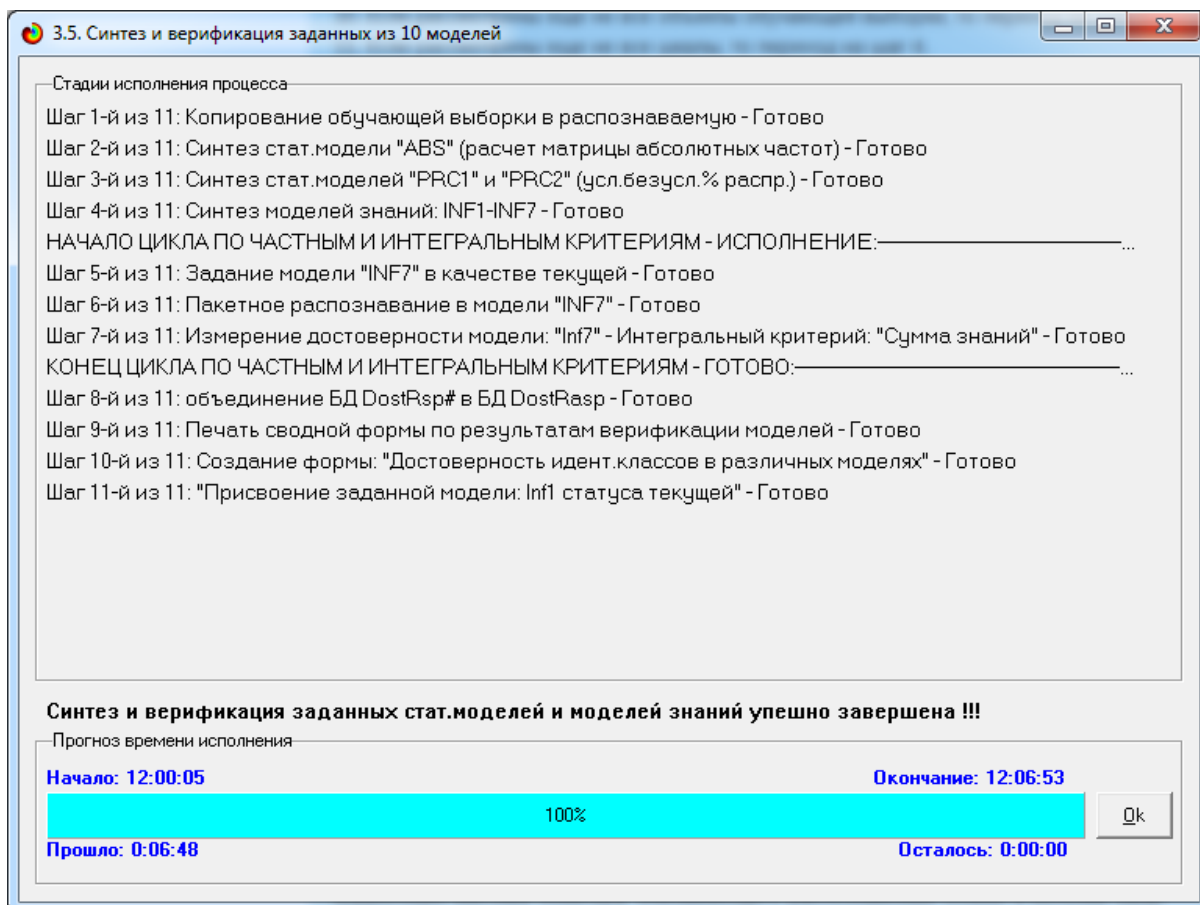


Рисунок 6. Отображение стадии исполнения режима 3.5.

Достоверность идентификации и не идентификации объектов обучающей выборки в созданных моделях приведена в экранной форме режима 4.1.3.6 (рисунок 7):

| Наименование модели и частного критерия | Интегральный критерий | Вероятность правильной идентификац... | Вероятность правильной не идентиф... | Средняя вероятно... результата | Дата получения результата | Время получения результ... |
|---|-----------------------------------|---------------------------------------|--------------------------------------|--------------------------------|---------------------------|----------------------------|
| 1. ABS - частный критерий: количество встреч сочетаний "клас... | Корреляция абс. частот с обр... | 100.000 | 0.058 | 50.029 | 25.07.2016 | 12:01:25 |
| 2. PRC1 - частный критерий: усл. вероятность н-го признака сред... | Корреляция усл.отн. частот с о... | 100.000 | 0.058 | 50.029 | 25.07.2016 | 12:02:02 |
| 3. PRC2 - частный критерий: условная вероятность н-го признака... | Корреляция усл.отн. частот с о... | 100.000 | 0.058 | 50.029 | 25.07.2016 | 12:02:39 |
| 4. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 90.112 | 79.241 | 84.677 | 25.07.2016 | 12:03:15 |
| 5. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Сумма знаний | 82.022 | 84.325 | 83.174 | 25.07.2016 | 12:03:15 |
| 6. INF3 - частный критерий: Хинкватрат, разности между фактич... | Семантический резонанс зна... | 90.169 | 79.098 | 84.633 | 25.07.2016 | 12:03:51 |
| 7. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 89.101 | 81.296 | 85.198 | 25.07.2016 | 12:05:04 |
| 8. INF5 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 92.640 | 66.270 | 79.455 | 25.07.2016 | 12:05:04 |
| 9. INF6 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 87.809 | 80.885 | 84.347 | 25.07.2016 | 12:05:40 |
| 10. INF7 - частный критерий: разн.усли. безуслов. вероятностей; ве... | Семантический резонанс зна... | 91.292 | 65.658 | 78.475 | 25.07.2016 | 12:05:40 |
| 1. INF1 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 90.225 | 72.915 | 81.570 | 25.07.2016 | 12:06:16 |
| 2. INF2 - частный критерий: количество знаний по А.Харкевичу; в... | Семантический резонанс зна... | 91.011 | 66.401 | 78.706 | 25.07.2016 | 12:06:16 |
| 3. INF3 - частный критерий: Хинкватрат, разности между фактич... | Семантический резонанс зна... | 89.213 | 72.689 | 80.951 | 25.07.2016 | 12:06:52 |
| 4. INF4 - частный критерий: ROI (Return On Investment); вероятно... | Семантический резонанс зна... | 88.202 | 65.564 | 76.883 | 25.07.2016 | 12:06:52 |

Помощь

Рисунок 7. Достоверность идентификации и не идентификации объектов обучающей выборки в созданных моделях

Для оценки достоверности моделей используется предложенная проф. Е.В. Луценко метрика, сходная с известным F-критерием (рисунок 8):

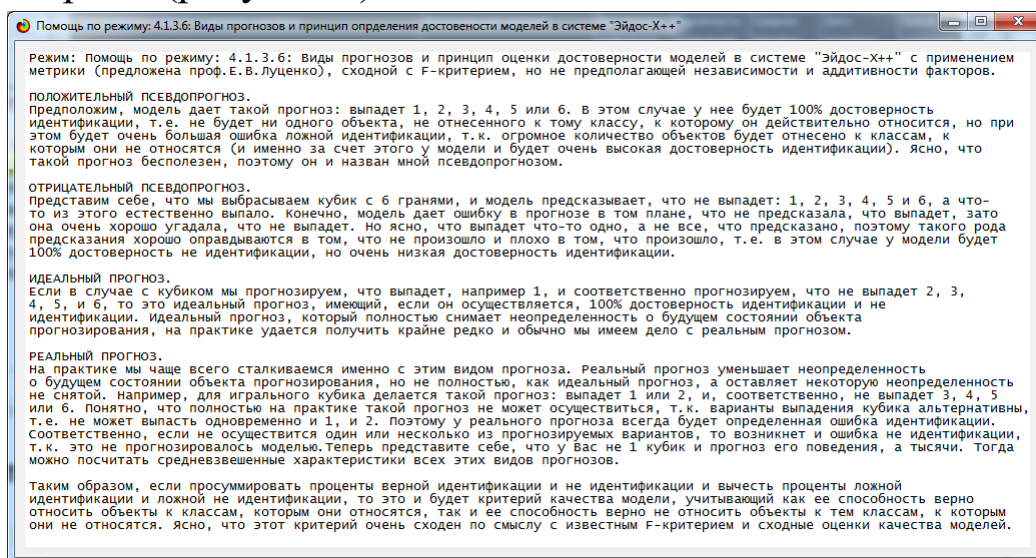


Рисунок 8. Оценка достоверности моделей в системе «Эйдос»

Ниже, на рисунках 9 приведены фрагменты моделей, отражающих взаимосвязи между структурой генома и фенотипическими признаками, и свойствами винограда, отличающихся частными критериями знаний [8, 9] (экранные формы получены в режиме 5.5, названия моделей приведены на экранных формах).

| Код признака | Наименование описательной шкалы и градации | 1. ТИП ЦВЕТКА (151)-1/2:3 | 2. ТИП ЦВЕТКА (151)-2/2:4 | 3. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ МАЛОЕ: 1/5-(27.00...112.000000... | 4. СРЕДНЯЯ МАССА ГРОЗДИ, Г-МАЛОЕ: 2/5-(112.0...173.000000... | 5. СРЕДНЯЯ МАССА ГРОЗДИ, Г-СРЕДН... 3/5-(173.0...220.000000... | 6. СРЕДНЯЯ МАССА ГРОЗДИ, Г-БОЛЬШ... 4/5-(220.0...280.000000... | 7. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ БОЛЬШОЕ: 5/5-(280.0...400.000000... | 8. УРОЖАЙ, Ц/ГА-ОЧЕ... МАЛОЕ: 1/5-(9.000...60.00000000 |
|--------------|---|---------------------------|---------------------------|--|--|--|--|--|--|
| 1 | VRZAG47-2-Очень мало: 1/5-(176.0000000, 180.000... | 26 | 5 | 6 | 8 | 7 | 6 | 4 | 5 |
| 2 | VRZAG47-2-Мало: 2/5-(180.0000000, 182.0000000) ... | 17 | 10 | 7 | 3 | 2 | 5 | 10 | 7 |
| 3 | VRZAG47-2-Среднее: 3/5-(182.0000000, 182.0000000... | | | | | | | | |
| 4 | VRZAG47-2-Большое: 4/5-(182.0000000, 186.0000000... | 16 | 3 | 2 | 6 | 3 | 5 | 3 | 2 |
| 5 | VRZAG47-2-Очень большое: 5/5-(186.0000000, 195.0... | 11 | 1 | 2 | 1 | 6 | 2 | 1 | 3 |
| 6 | VRZAG47-1-Очень мало: 1/5-(180.0000000, 184.000... | 15 | 4 | 4 | 5 | 4 | 3 | 3 | 6 |
| 7 | VRZAG47-1-Мало: 2/5-(184.0000000, 186.0000000) ... | 16 | 2 | 3 | 3 | 1 | 3 | 8 | 2 |
| 8 | VRZAG47-1-Среднее: 3/5-(186.0000000, 190.0000000... | 14 | 5 | 5 | 6 | 6 | 1 | 1 | 4 |
| 9 | VRZAG47-1-Большое: 4/5-(190.0000000, 195.0000000... | 25 | 8 | 5 | 4 | 7 | 11 | 6 | 5 |
| 10 | VRZAG47-1-Очень большое: 5/5-(195.0000000, 195.0... | | | | | | | | |
| 11 | WS2-2-Очень мало: 1/5-(123.0000000, 133.0000000... | 17 | 7 | 6 | 6 | 4 | 6 | 2 | 5 |
| 12 | WS2-2-Мало: 2/5-(133.0000000, 135.0000000) | 29 | 6 | 4 | 4 | 8 | 9 | 10 | 4 |
| 13 | WS2-2-Среднее: 3/5-(135.0000000, 135.0000000) ... | | | | | | | | |
| 14 | WS2-2-Большое: 4/5-(135.0000000, 143.0000000) ... | 16 | 4 | 5 | 7 | 3 | 2 | 3 | 5 |
| 15 | WS2-2-Очень большое: 5/5-(143.0000000, 149.0000... | 7 | 2 | 2 | 1 | 2 | 1 | 3 | 3 |

5.5. Модель: "3. PRC2 - частный критерий: условная вероятность i-го признака у объектов j-го класса"

| Код признака | Наименование описательной шкалы и градации | 1. ТИП ЦВЕТКА (151)-1/2:3 | 2. ТИП ЦВЕТКА (151)-2/2:4 | 3. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ МАЛОЕ: 1/5-(27.00... 112.000000... | 4. СРЕДНЯЯ МАССА ГРОЗДИ, Г-МАЛОЕ: 2/5-(112.0... 173.000000... | 5. СРЕДНЯЯ МАССА ГРОЗДИ, Г-СРЕДН... 3/5-(173.0... 220.000000... | 6. СРЕДНЯЯ МАССА ГРОЗДИ, Г-БОЛЬШ... 4/5-(220.0... 280.000000... | 7. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ БОЛЬШОЕ: 5/5-(280.0... 400.000000... | 8. УРОЖАЙ... Ц/ГА-ОЧЕ... МАЛОЕ: 1/5-(9.000... 60.00000000) | 9. УРОЖАЙ... Ц/ГА-ОЧЕ... БОЛЬШОЕ: 2/5-(9.000... 95.00000000) |
|--------------|---|---------------------------|---------------------------|---|---|---|---|---|--|--|
| 1 | VRZAG47-2-Очень малое: 1/5-(176.0000000, 180... | 37.143 | 26.316 | 35.294 | 44.444 | 38.889 | 33.333 | 22.222 | 29.412 | |
| 2 | VRZAG47-2-Малое: 2/5-(180.0000000, 182.00000... | 24.286 | 52.632 | 41.176 | 16.667 | 11.111 | 27.778 | 55.556 | 41.176 | |
| 3 | VRZAG47-2-Среднее: 3/5-(182.0000000, 182.000... | | | | | | | | | |
| 4 | VRZAG47-2-Большое: 4/5-(182.0000000, 186.000... | 22.857 | 15.789 | 11.765 | 33.333 | 16.667 | 27.778 | 16.667 | 11.765 | |
| 5 | VRZAG47-2-Очень большое: 5/5-(186.0000000, 1... | 15.714 | 5.263 | 11.765 | 5.556 | 33.333 | 11.111 | 5.556 | 17.647 | |
| 6 | VRZAG47-1-Очень малое: 1/5-(180.0000000, 184... | 21.429 | 21.053 | 23.529 | 27.778 | 22.222 | 16.667 | 16.667 | 35.294 | |
| 7 | VRZAG47-1-Малое: 2/5-(184.0000000, 186.00000... | 22.857 | 10.526 | 17.647 | 16.667 | 5.556 | 16.667 | 44.444 | 11.765 | |
| 8 | VRZAG47-1-Среднее: 3/5-(186.0000000, 190.000... | 20.000 | 26.316 | 29.412 | 33.333 | 33.333 | 5.556 | 5.556 | 23.529 | |
| 9 | VRZAG47-1-Большое: 4/5-(190.0000000, 195.000... | 35.714 | 42.105 | 29.412 | 22.222 | 38.889 | 61.111 | 33.333 | 29.412 | |
| 10 | VRZAG47-1-Очень большое: 5/5-(195.0000000, 1... | | | | | | | | | |
| 11 | WS2-2-Очень малое: 1/5-(123.0000000, 133.000... | 24.286 | 36.842 | 35.294 | 33.333 | 22.222 | 33.333 | 11.111 | 29.412 | |
| 12 | WS2-2-Малое: 2/5-(133.0000000, 135.0000000) ... | 41.429 | 31.579 | 23.529 | 22.222 | 44.444 | 50.000 | 55.556 | 23.529 | |
| 13 | WS2-2-Среднее: 3/5-(135.0000000, 135.0000000... | | | | | | | | | |
| 14 | WS2-2-Большое: 4/5-(135.0000000, 143.000000... | 22.857 | 21.053 | 29.412 | 38.889 | 16.667 | 11.111 | 16.667 | 29.412 | |
| 15 | WS2-2-Очень большое: 5/5-(143.0000000, 149.0... | 10.000 | 10.526 | 11.765 | 5.556 | 11.111 | 5.556 | 16.667 | 17.647 | |

Помощь MS Excel MS Word

5.5. Модель: "4. INF1 - частный критерий: количество знаний по А.Харкевичу; вероятности из PRC1"

| Код признака | Наименование описательной шкалы и градации | 1. ТИП ЦВЕТКА (151)-1/2:3 | 2. ТИП ЦВЕТКА (151)-2/2:4 | 3. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ МАЛОЕ: 1/5-(27.00... 112.000000... | 4. СРЕДНЯЯ МАССА ГРОЗДИ, Г-МАЛОЕ: 2/5-(112.0... 173.000000... | 5. СРЕДНЯЯ МАССА ГРОЗДИ, Г-СРЕДН... 3/5-(173.0... 220.000000... | 6. СРЕДНЯЯ МАССА ГРОЗДИ, Г-БОЛЬШ... 4/5-(220.0... 280.000000... | 7. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ БОЛЬШОЕ: 5/5-(280.0... 400.000000... | 8. УРОЖАЙ... Ц/ГА-ОЧЕ... МАЛОЕ: 1/5-(9.000... 60.00000000) | 9. УРОЖАЙ... Ц/ГА-ОЧЕ... БОЛЬШОЕ: 2/5-(9.000... 95.00000000) |
|--------------|---|---------------------------|---------------------------|---|---|---|---|---|--|--|
| 1 | VRZAG47-2-Очень малое: 1/5-(176.0000000, 180... | 0.041 | -0.179 | 0.006 | 0.138 | 0.089 | -0.023 | -0.285 | -0.074 | |
| 2 | VRZAG47-2-Малое: 2/5-(180.0000000, 182.00000... | -0.129 | 0.317 | 0.180 | -0.364 | -0.575 | -0.049 | 0.343 | 0.209 | |
| 3 | VRZAG47-2-Среднее: 3/5-(182.0000000, 182.000... | | | | | | | | | |
| 4 | VRZAG47-2-Большое: 4/5-(182.0000000, 186.000... | 0.044 | -0.191 | -0.357 | 0.258 | -0.124 | 0.160 | -0.165 | -0.328 | |
| 5 | VRZAG47-2-Очень большое: 5/5-(186.0000000, 1... | 0.094 | -0.572 | -0.083 | -0.536 | 0.563 | -0.112 | -0.545 | 0.187 | |
| 6 | VRZAG47-1-Очень малое: 1/5-(180.0000000, 184... | 0.005 | -0.020 | 0.056 | 0.149 | 0.047 | -0.144 | -0.165 | 0.327 | |
| 7 | VRZAG47-1-Малое: 2/5-(184.0000000, 186.00000... | 0.076 | -0.401 | -0.083 | -0.123 | -0.746 | -0.112 | 0.452 | -0.296 | |
| 8 | VRZAG47-1-Среднее: 3/5-(186.0000000, 190.000... | -0.036 | 0.113 | 0.189 | 0.258 | 0.289 | -0.798 | -0.819 | 0.085 | |
| 9 | VRZAG47-1-Большое: 4/5-(190.0000000, 195.000... | -0.019 | 0.064 | -0.140 | -0.313 | 0.052 | 0.301 | -0.081 | -0.111 | |
| 10 | VRZAG47-1-Очень большое: 5/5-(195.0000000, 1... | | | | | | | | | |
| 11 | WS2-2-Очень малое: 1/5-(123.0000000, 133.000... | -0.059 | 0.174 | 0.159 | 0.119 | -0.092 | 0.130 | -0.545 | 0.079 | |
| 12 | WS2-2-Малое: 2/5-(133.0000000, 135.0000000) ... | 0.034 | -0.142 | -0.308 | -0.348 | 0.096 | 0.147 | 0.189 | -0.279 | |
| 13 | WS2-2-Среднее: 3/5-(135.0000000, 135.0000000... | | | | | | | | | |
| 14 | WS2-2-Большое: 4/5-(135.0000000, 143.000000... | 0.013 | -0.050 | 0.159 | 0.319 | -0.154 | -0.416 | -0.195 | 0.187 | |
| 15 | WS2-2-Очень большое: 5/5-(143.0000000, 149.0... | -0.003 | 0.012 | 0.089 | -0.364 | 0.080 | -0.353 | 0.280 | 0.359 | |

Помощь MS Excel MS Word

5.5. Модель: "6. INF3 - частный критерий: Хи-квадрат, разности между фактическими и ожидаемыми абсчастотами"

| Код признака | Наименование описательной шкалы и градации | 1. ТИП ЦВЕТКА (151)1/2-3 | 2. ТИП ЦВЕТКА (151)2/2-4 | 3. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ МАЛОЕ: 1/5(127.00...112.000000... | 4. СРЕДНЯЯ МАССА ГРОЗДИ, Г-МАЛОЕ: 2/5(112.0...173.000000... | 5. СРЕДНЯЯ МАССА ГРОЗДИ, Г-СРЕДН... 3/5(173.0...220.000000... | 6. СРЕДНЯЯ МАССА ГРОЗДИ, Г-БОЛЬШ... 4/5(220.0...280.000000... | 7. СРЕДНЯЯ МАССА ГРОЗДИ, Г-ОЧЕНЬ БОЛЬШОЕ: 5/5(280.0...400.000000... | 8. УРОЖАЙ... Ц/ТА-ОЧЕ... МАЛОЕ: 1/5(9.000...60.00000000) | 9. Ц/ТА-ОЧЕ... МАЛОЕ: 2/5(9.000...60.00000000) |
|--------------|--|--------------------------|--------------------------|--|---|---|---|---|--|--|
| 1 | VRZAG47-2-Очень мало: 1/5(176.0000000, 180... | 1.748 | -1.748 | 0.062 | 1.650 | 0.974 | -0.232 | -2.453 | -0.658 | |
| 2 | VRZAG47-2-Мало: 2/5(180.0000000, 182.000000... | -4.123 | 4.123 | 1.828 | -2.531 | -3.249 | -0.428 | 4.379 | 2.072 | |
| 3 | VRZAG47-2-Среднее: 3/5(182.0000000, 182.000... | | | | | | | | | |
| 4 | VRZAG47-2-Большое: 4/5(182.0000000, 186.000... | 1.136 | -1.136 | -1.639 | 2.108 | -0.693 | 1.180 | -0.955 | -1.468 | |
| 5 | VRZAG47-2-Очень большое: 5/5(186.0000000, 1... | 1.612 | -1.612 | -0.298 | -1.458 | 3.667 | -0.413 | -1.498 | 0.810 | |
| 6 | VRZAG47-1-Очень мало: 1/5(180.0000000, 184... | 0.136 | -0.136 | 0.361 | 1.108 | 0.307 | -0.820 | -0.955 | 2.532 | |
| 7 | VRZAG47-1-Мало: 2/5(184.0000000, 186.000000... | 1.918 | -1.918 | -0.448 | -0.687 | -2.499 | -0.619 | 4.253 | -1.285 | |
| 8 | VRZAG47-1-Среднее: 3/5(186.0000000, 190.000... | -0.864 | 0.864 | 1.361 | 2.108 | 2.307 | -2.820 | -2.955 | 0.532 | |
| 9 | VRZAG47-1-Большое: 4/5(190.0000000, 195.000... | -0.817 | 0.817 | -1.321 | -2.760 | 0.585 | 4.365 | -0.870 | -1.023 | |
| 10 | VRZAG47-1-Очень большое: 5/5(195.0000000, 1... | | | | | | | | | |
| 11 | VS2-2-Очень мало: 1/5(123.0000000, 133.0000... | -1.776 | 1.776 | 1.403 | 1.084 | -0.665 | 1.175 | -2.996 | 0.620 | |
| 12 | VS2-2-Мало: 2/5(133.0000000, 135.0000000) ... | 1.619 | -1.619 | -2.704 | -3.170 | 1.196 | 1.963 | 2.714 | -2.388 | |
| 13 | VS2-2-Среднее: 3/5(135.0000000, 135.0000000... | | | | | | | | | |
| 14 | VS2-2-Большое: 4/5(135.0000000, 143.000000... | 0.354 | -0.354 | 1.169 | 2.903 | -0.888 | -2.021 | -1.163 | 1.350 | |
| 15 | VS2-2-Очень большое: 5/5(143.0000000, 149.0... | -0.041 | 0.041 | 0.276 | -0.844 | 0.250 | -0.809 | 1.126 | 1.357 | |

Рисунок 9. Фрагменты моделей, отражающих взаимосвязи между структурой генома и фенотипическими свойствами/признаками винограда

4.4. Решение задачи системной идентификации фенотипических признаков и свойств по наличию тех или иных генов

Смысл этой задачи состоит в том, чтобы, используя отраженные в моделях знания о силе и направлении влияния различных структур генома (степень выраженности генов на спиральных ДНК) на степень выраженности различных фенотипических признаков и свойств по описанию структур генома, определить, какими фенотипическими признаками и свойствами будет обладать виноград с таким геномом.

В АСК-анализе и системе «Эйдос» эта задача решается двумя способами:

- 1) считается, что объект тем в большей степени относится к некоторому классу, чем больше информации о принадлежности к нему содержится в его признаках;
- 2) считается, что объект тем в большей степени относится к некоторому классу, чем больше его профиль по степени

выраженности признаков сходен с профилем класса, в котором отражена степень характерности этих признаков для класса.

Первый способ соответствует применению интегрального критерия: «Сумма знаний», а второй – «Резонанс знаний».

Рассмотрим интегральные критерии знаний, используемые в настоящее время в АСК-анализе и системе «Эйдос-Х++» для верификации моделей и решения задач идентификации и прогнозирования, а затем приведем некоторые выходные формы с результатами распознавания. Все их не приведем, т.к. их довольно много (рисунок 10):

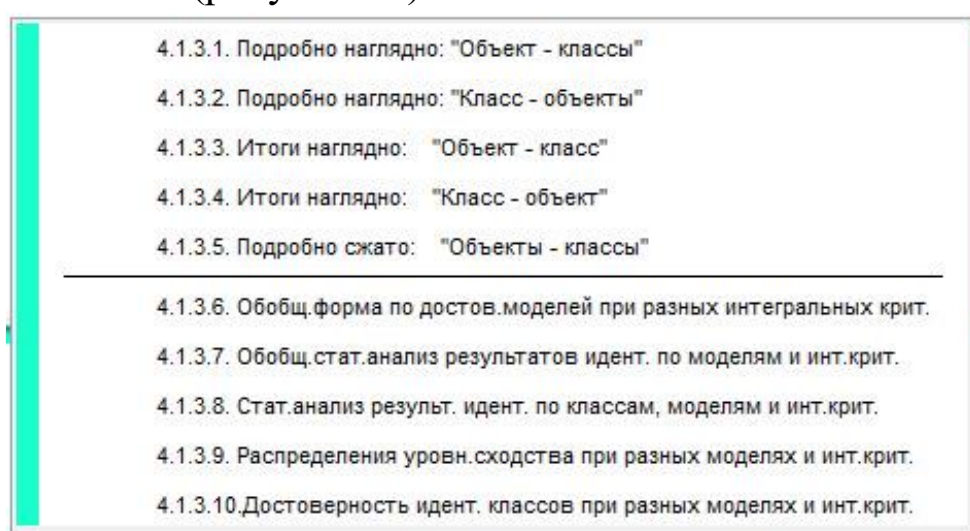


Рисунок 10. Формы с результатами распознавания системы «Эйдос»

4.4.1. Интегральный критерий – «Сумма знаний»

1-й интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний и имеет вид:

$$I_j = (\vec{I}_{ij}, \vec{L}_i).$$

В выражении круглыми скобками обозначено скалярное произведение. В координатной форме это выражение имеет вид:

$$I_j = \sum_{i=1}^M I_{ij} L_i,$$

где: M – количество градаций описательных шкал (признаков);

$\vec{I}_{ij} = \{I_{ij}\}$ – вектор состояния j -го класса;

$\vec{L}_i = \{L_i\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив-локатор), т.е.:

$$\vec{L}_i = \begin{cases} 1, & \text{если } i\text{-й фактор действует;} \\ n, & \text{где } n > 0, \text{ если } i\text{-й фактор действует с истинностью } n; \\ 0, & \text{если } i\text{-й фактор не действует.} \end{cases}$$

В текущей версии системы «Эйдос-Х++» значения координат вектора состояния распознаваемого объекта принимались равными либо 0, если признака нет, или n , если он присутствует у объекта с интенсивностью n , т.е. представлен n раз (например, буква «о» в слове «молоко» представлена 3 раза, а буква «м» - один раз).

4.4.2. Интегральный критерий – «Резонанс знаний»

2-й интегральный критерий «Семантический резонанс знаний» представляет собой *нормированное* суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния.

Интегральный критерий представляет собой аддитивную функцию от частных критериев знаний и имеет вид:

$$I_j = \frac{1}{\sigma_j \sigma_l M} \sum_{i=1}^M (I_{ij} - \bar{I}_j)(L_i - \bar{L}),$$

где:

M – количество градаций описательных шкал (признаков);

\bar{I}_j – средняя информативность по вектору класса;

\bar{L} – среднее по вектору объекта;

σ_j – среднее квадратичное отклонение частных критериев знаний вектора класса;

σ_l – среднее квадратичное отклонение по вектору распознаваемого объекта.

Приведенное выражение для интегрального критерия «Семантический резонанс знаний» получается непосредственно из выражения для критерия «Сумма знаний» после замены координат перемножаемых векторов их стандартизированными значениями:

$$I_{ij} \rightarrow \frac{I_{ij} - \bar{I}_j}{\sigma_j}, \quad L_i \rightarrow \frac{L_i - \bar{L}}{\sigma_l}.$$

Свое наименование интегральный критерий сходства «Семантический резонанс знаний» получил потому, что по своей математической форме является корреляцией двух векторов: состояния j -го класса и состояния распознаваемого объекта.

4.4.3. Идентификация объекта с классами

Экранная форма, в которой для каждого объекта обучающей выборки приводятся все классы в порядке убывания релевантности, приведена на рисунке 11:

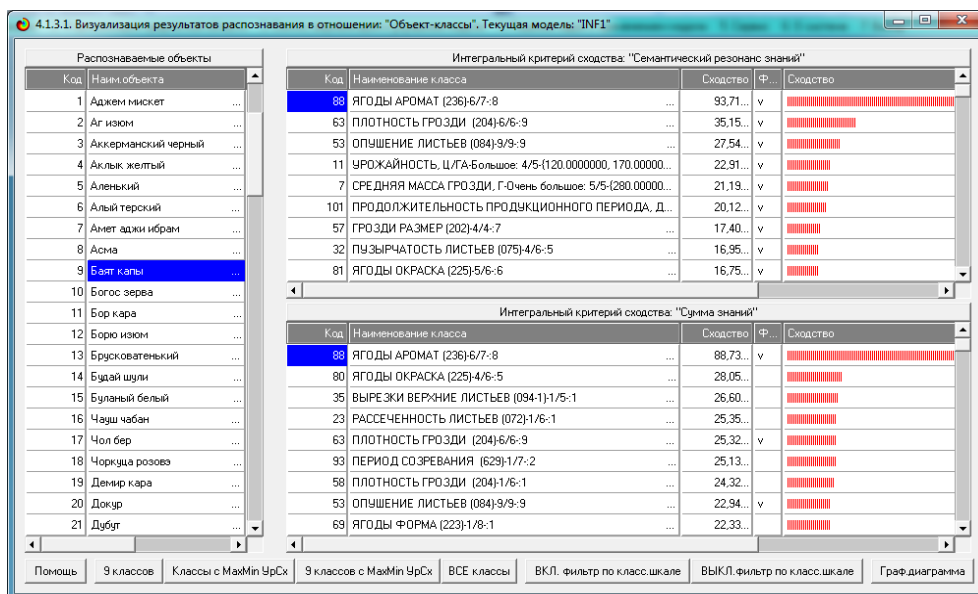


Рисунок 11. Результаты решения задачи идентификации фенотипических признаков и свойств сорта винограда по структуре его генома

4.4.4. Степень сходства объектов с классом

Экранная форма, в которой для каждого объекта обучающей выборки приводятся все классы в порядке убывания релевантности, приведена на рисунке 12:

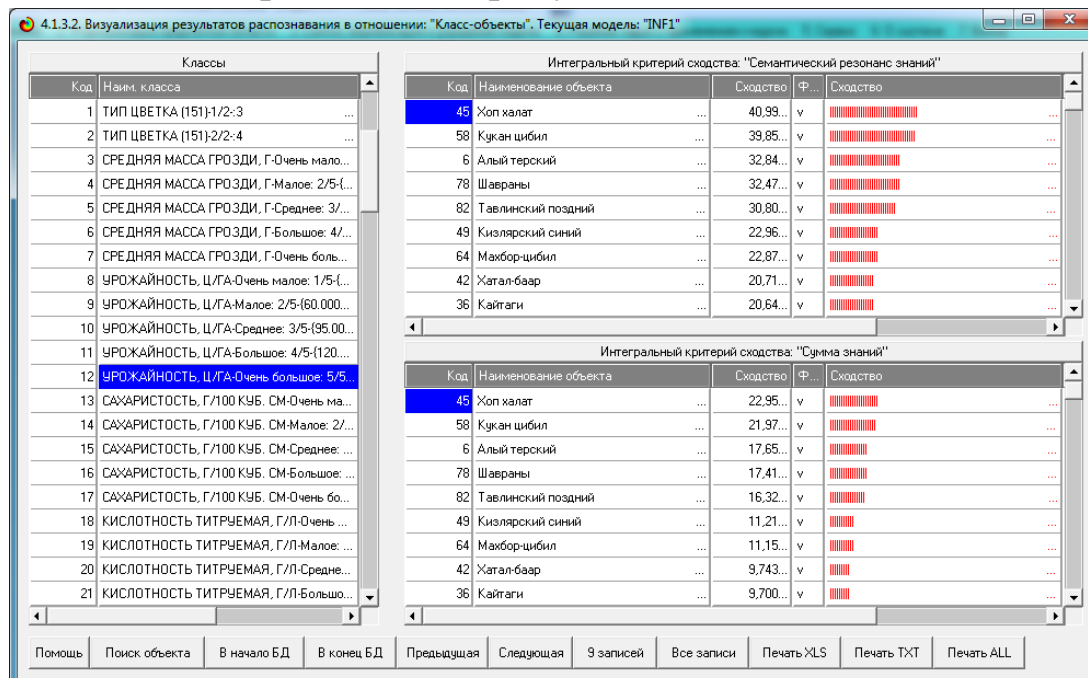


Рисунок 12. Результаты решения задачи идентификации сортов винограда с заданными фенотипическими признаками и свойствами по структуре их генома

4.5. Автоматизированный SWOT-анализ системы детерминации фенотипических признаков и свойств структурой генома

На экранных формах режима 4.4.8 системы «Эйдос», приведенных на рисунке 13, пользователь может выбрать любое фенотипическое свойство или признак и модель, в которой отображать систему детерминации этого фенотипического свойства/признака структурой генома и степенью выраженности различных генов в двух спиральях ДНК [18]:

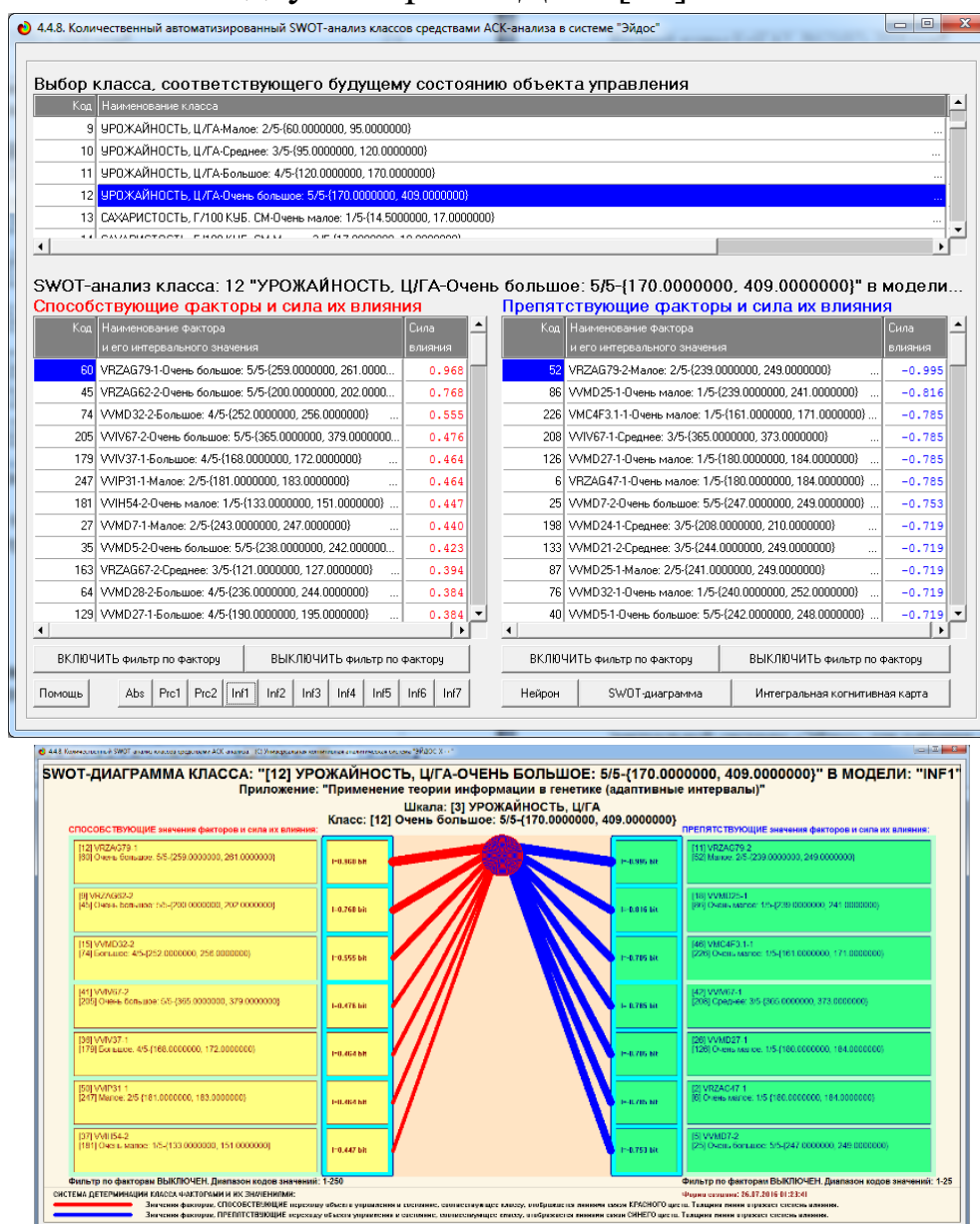
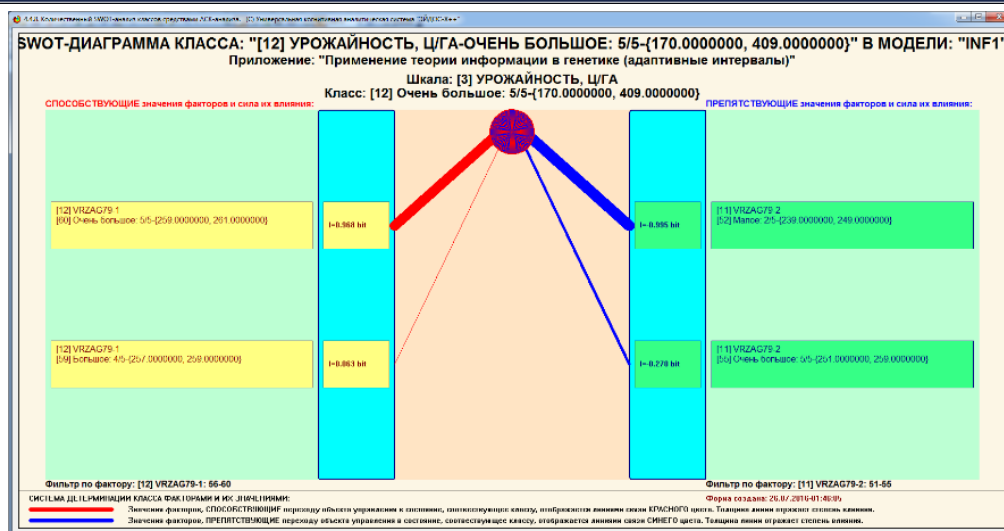
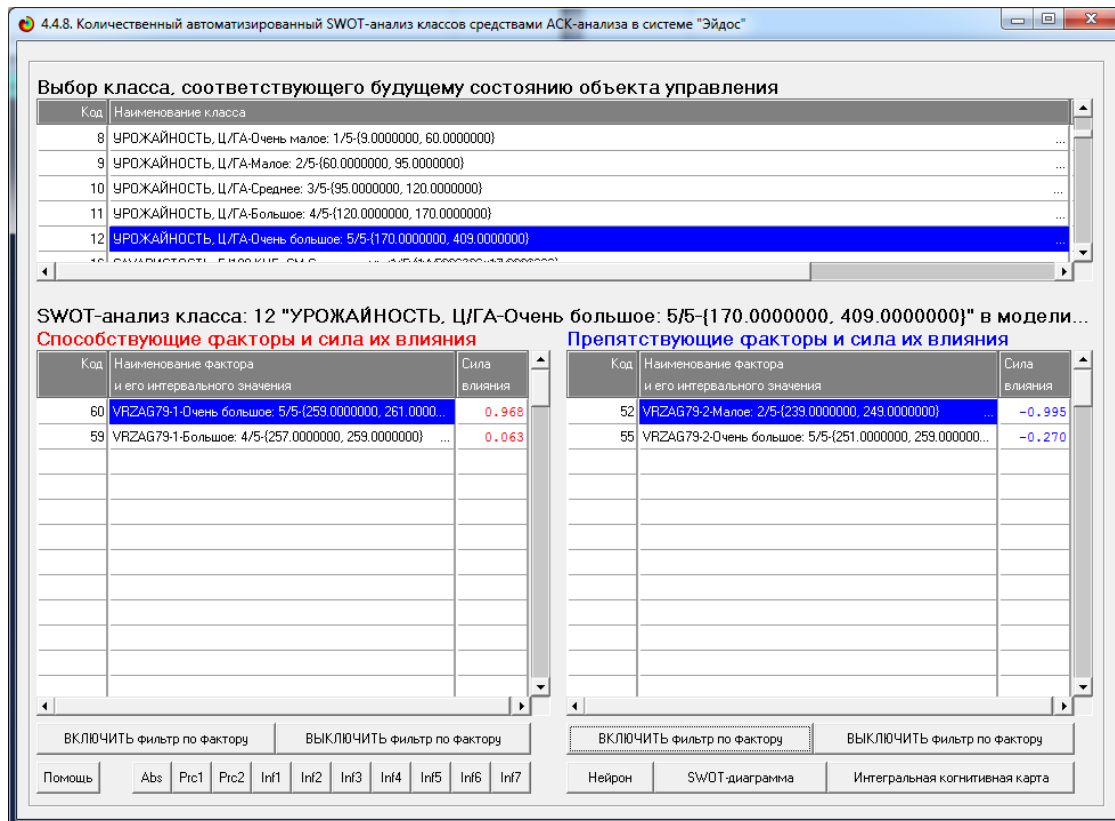


Рисунок 13. Экранные формы автоматизированного SWOT-анализа, показывающие, какие гены способствуют, а какие препятствуют проявлению заданного фенотипического признака/свойства в модели INF1

В верхней экранной форме, приведенной на рисунке 13, пользователь может выбрать любой ген и включить фильтр по нему. Тогда на экранных формах будет отражено влияние степени выраженности именно этого гена на заданное фенотипическое свойство (рисунок 14):



Рисунки 14. Экранные формы автоматизированного SWOT-анализа, показывающие, что высокая степень выраженности гена VRZAG79-1 способствует, а низкая VRZAG79-2 препятствует проявлению фенотипического свойства: «Очень большая урожайность» в модели INF1

4.6. Автоматизированный SWOT-анализ силы и направления влияния конкретного гена на фенотипические свойства/признаки

На экранных формах режима 4.4.9 системы «Эйдос», приведенных на рисунке 15, пользователь может выбрать любой ген и степень его выраженности в 1-й или 2-й аллели и модель, в которой отображать - какое влияние он оказывает на фенотипические свойства/признаки:

4.4.9 Количественный автоматизированный SWOT-анализ значений факторов средствами АСК-анализа в системе "Эйдос"

Выбор значения фактора, оказывающего влияние на переход объекта управления в будущее состояния

| Код | Наименование значения фактора |
|-----|---|
| 57 | VRZAG79-1-Малое: 2/5-(251.0000000, 251.0000000) |
| 58 | VRZAG79-1-Среднее: 3/5-(251.0000000, 257.0000000) |
| 59 | VRZAG79-1-Большое: 4/5-(257.0000000, 259.0000000) |
| 60 | VRZAG79-1-Очень большое: 5/5-(259.0000000, 261.0000000) |
| 61 | WMD29-2-Очень малое: 1/5-(218.0000000, 234.0000000) |

SWOT-анализ значения фактора: 60 "VRZAG79-1-Очень большое: 5/5-(259.0000000, 261.0000000)" в моде...

СПОСОБСТВУЕТ:

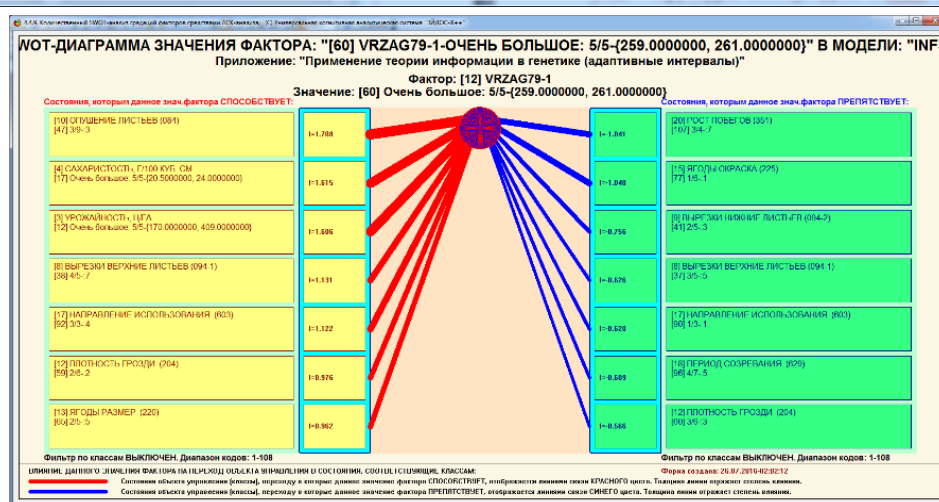
| Код | Состояния объекта управления, переходу в которые данное значение фактора СПОСОБСТВУЕТ | Сила влияния |
|-----|---|--------------|
| 47 | ОПУШЕНИЕ ЛИСТЬЕВ (084) 3/9-3 | 1.708 |
| 17 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Очень большое: 5/5-(2... | 1.615 |
| 12 | УРОЖАЙНОСТЬ, Ц/ГА-Очень большое: 5/5-(170.000000... | 1.606 |
| 38 | ВЫРЕЗКИ ВЕРХНИЕ ЛИСТЬЕВ (094-1) 4/5-7 | 1.131 |
| 92 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603) 3/3-4 | 1.122 |
| 59 | ПЛОТНОСТЬ ГРОЗДИ (204) 2/6-2 | 0.976 |
| 65 | ЯГОДЫ РАЗМЕР (220) 2/5-5 | 0.962 |
| 106 | РОСТ ПЕГОГОВ (351) 2/4-6 | 0.838 |
| 83 | ЯГОДЫ АРОМАТ (236) 1/7-1 | 0.809 |
| 56 | ГРОЗДИ РАЗМЕР (202) 3/4-6 | 0.799 |
| 103 | ПРОДОЛЖИТЕЛЬНОСТЬ ПРОДУКЦИОННОГО ПЕРИО... | 0.698 |
| 82 | ЯГОДЫ ОКРАСКА (225) 6/6-7 | 0.661 |

ПРЕПЯТСТВУЕТ:

| Код | Состояния объекта управления, переходу в которые данное значение фактора ПРЕПЯТСТВУЕТ | Сила влияния |
|-----|---|--------------|
| 107 | РОСТ ПЕГОГОВ (351) 3/4-7 | -1.041 |
| 77 | ЯГОДЫ ОКРАСКА (225) 1/6-1 | -1.040 |
| 41 | ВЫРЕЗКИ НИЖНИЕ ЛИСТЬЕВ (094-2) 2/5-3 | -0.756 |
| 37 | ВЫРЕЗКИ ВЕРХНИЕ ЛИСТЬЕВ (094-1) 3/5-5 | -0.626 |
| 90 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603) 1/3-1 | -0.620 |
| 96 | ПЕРИОД СОЗРЕВАНИЯ (629) 4/7-5 | -0.609 |
| 60 | ПЛОТНОСТЬ ГРОЗДИ (204) 3/6-3 | -0.566 |
| 45 | ОПУШЕНИЕ ЛИСТЬЕВ (084) 1/9-1 | -0.552 |
| 14 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Малое: 2/5-(17.0000000, ... | -0.551 |
| 85 | ЯГОДЫ АРОМАТ (236) 3/7-4 | -0.524 |
| 67 | ЯГОДЫ РАЗМЕР (220) 4/5-7 | -0.522 |
| 91 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603) 2/3-3 | -0.502 |

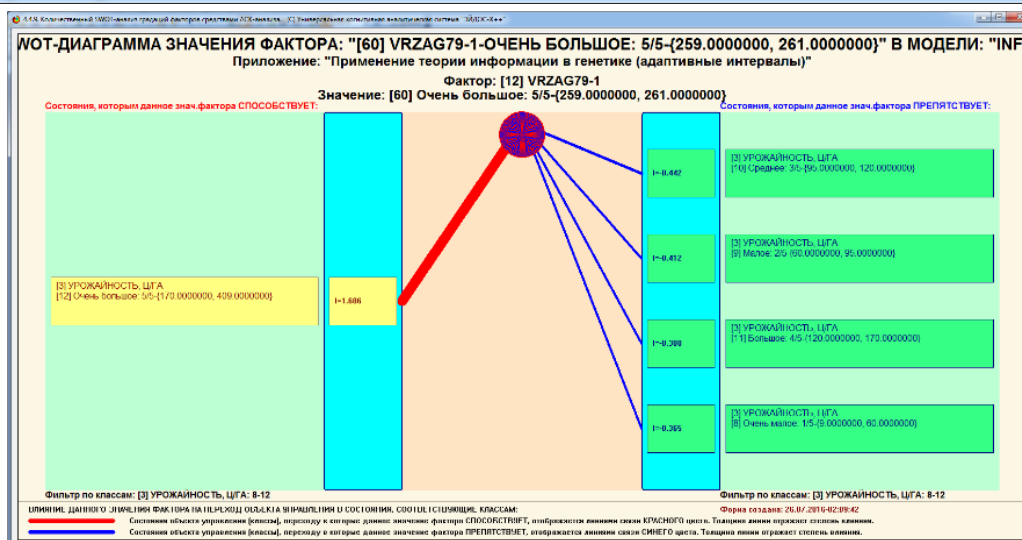
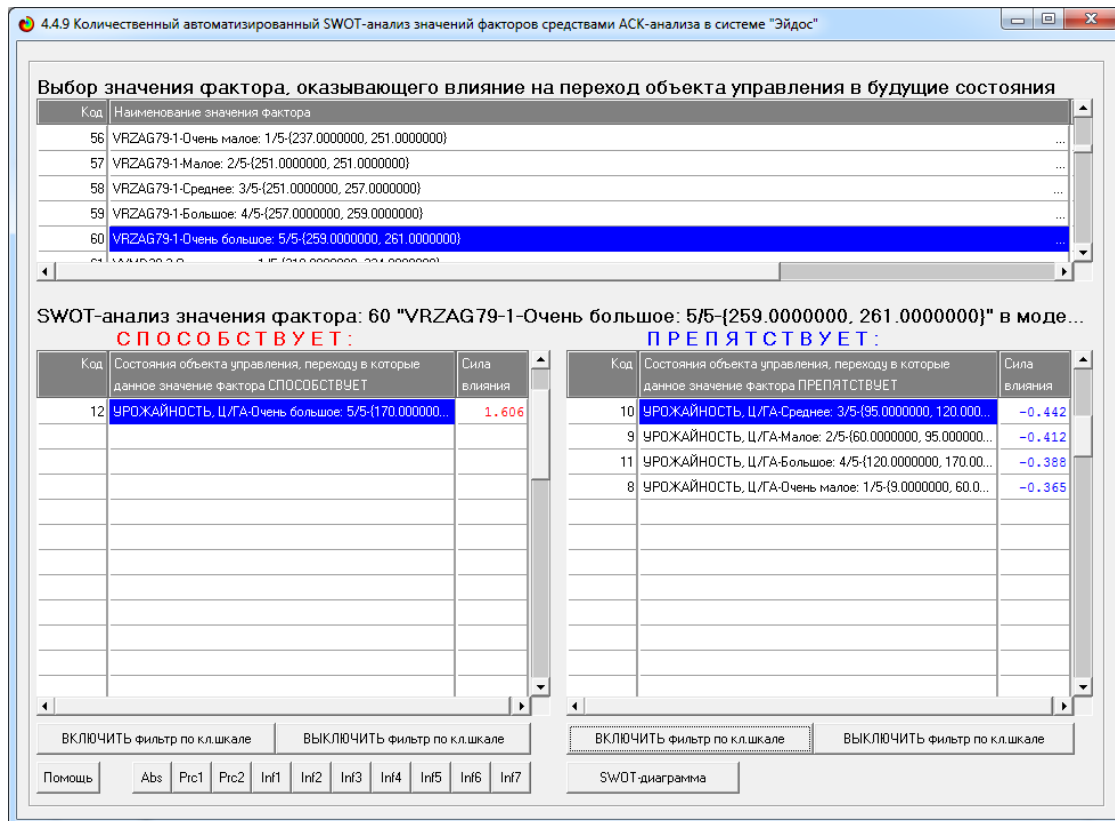
ВКЛЮЧИТЬ фильтр по кликалке ВЫКЛЮЧИТЬ фильтр по кликалке ВКЛЮЧИТЬ фильтр по кликалке ВЫКЛЮЧИТЬ фильтр по кликалке

Помощь Abs Pct1 Pct2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 SWOT-диаграмма



Рисунки 15. Экранные формы автоматизированного SWOT-анализа, показывающие, какое влияние заданный ген оказывает на степень выраженности различных фенотипических свойств/признаков в модели INF3

В верхней экранной форме, приведенной на рисунке 16, пользователь может выбрать любое фенотипическое свойство и включить фильтр по нему. Тогда на экранных формах будет отражено влияние заданной степени выраженности заданного гена на это фенотипическое свойство (рисунок 17):



Рисунки 16. Экранные формы автоматизированного SWOT-анализа, показывающие, что высокая степень выраженности гена VRZAG79-1 способствует очень высокой урожайности и препятствует всем остальным степеням выраженности урожайности в модели INF3

Таким образом, используя автоматизированный когнитивный SWOT-анализ генов и фенотипических свойств/признаков, мы можем получить информацию о том, какими генами детерминируется заданное фенотипическое свойство/признак и наоборот, какие фенотипические признаки/свойства обуславливаются заданным геном. Обладание этими знаниями вселяет надежду на возможность получения ответа на вопрос о том, *каким должен быть ген, чтобы сорт обладал заданными желательными фенотипическими свойствами.*

Однако здесь есть по меньшей мере одна проблема, которая состоит в том, что некоторые фенотипические свойства/признаки детерминируются наличием или сильной выраженностью определенных генов, а другие либо их полным отсутствием, либо слабой выраженностью. Ясно, что одновременно ген не может и присутствовать, и отсутствовать, быть и сильно, и слабо выраженным.

Из этого следует, что некоторые фенотипические свойства/признаки невозможно одновременно получить в одном сорте, а другие вполне возможно. Например, ягода не может быть и сладкой, и кислой, и темной, и светлой. Может быть в этом и состоит оправдание существования многих сортов. Чтобы узнать, конкретно какие свойства совместимы в одном сорте, а какие нет, необходимо провести сравнение различных фенотипических свойств по их системе детерминации генами.

4.7. Количественное определение сходства-различия различных фенотипических свойств по их системе детерминации генами. Совместимые и несовместимые фенотипические свойства

Эта задача решается в режиме 4.2.2 системы «Эйдос», в котором сначала рассчитываются матрицы сходства классов (рисунки 17 и 18), а затем они просто визуализируются в форме когнитивных диаграмм или, кроме того, на их основе проводится

когнитивный кластерный анализ [19]¹⁵. Результаты кластерного анализа классов приведены на табличной экранной форме режима 4.2.2.2 (рисунок 19) и в графической форме когнитивной диаграммы (рисунки 20 и 21).

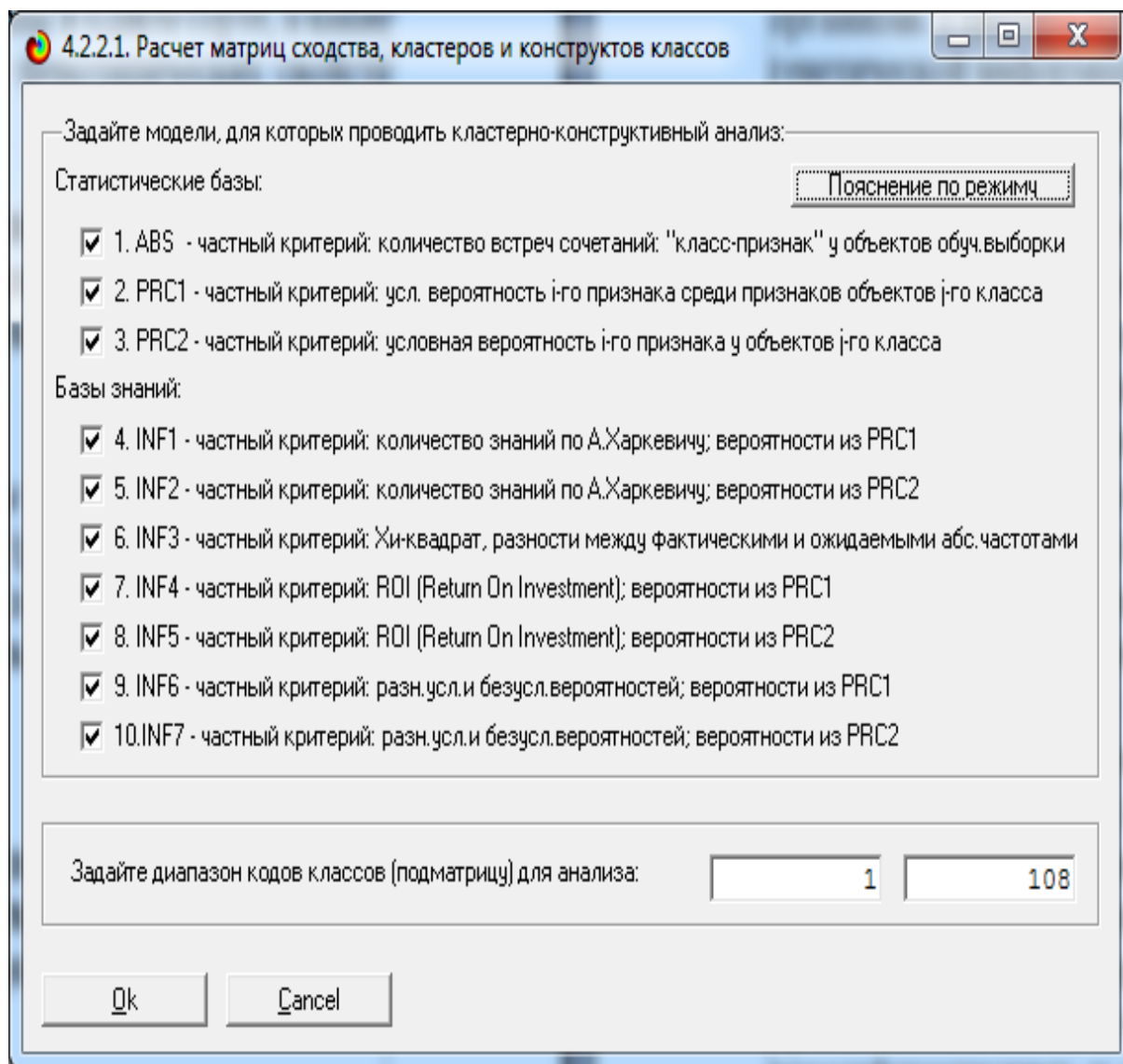


Рисунок 17. Форма задания параметров формирования матрицы сходства классов

¹⁵ Он был реализован в DOS-версии системы «Эйдос», а в новой версии его еще нет.

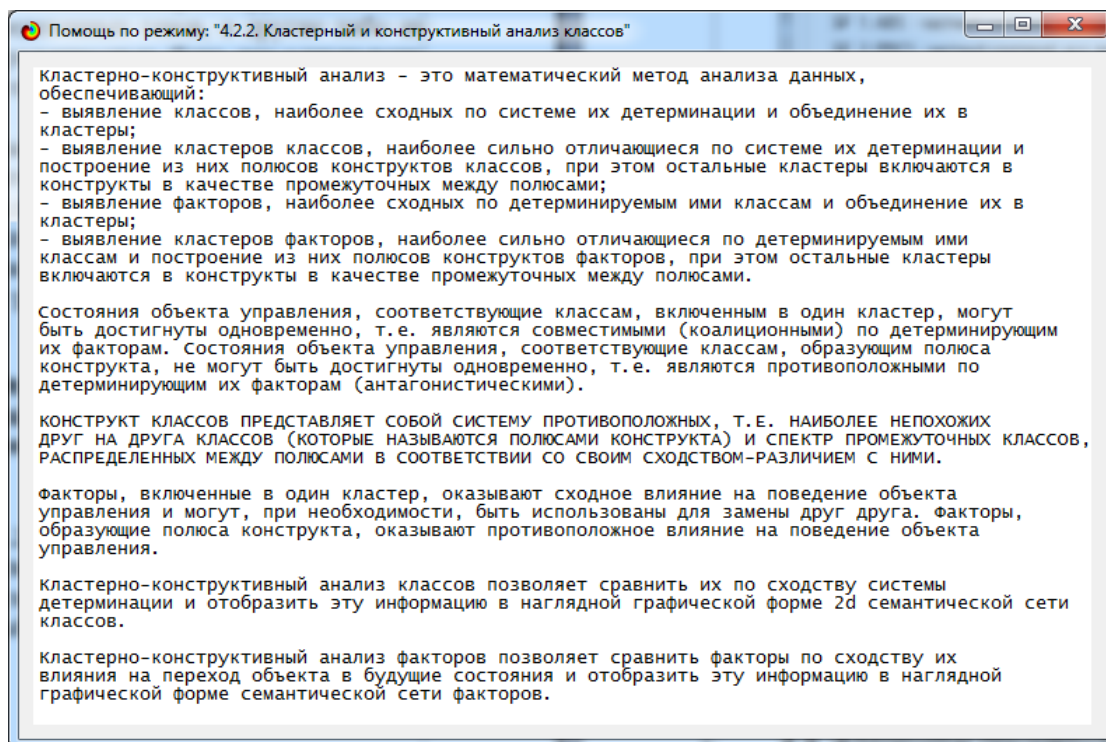


Рисунок 18. Help режима кластерного анализа классов системы «Эйдос»

4.2.2.2. Результаты кластерно-конструктивного анализа классов

Конструктор класса: 12 "УРОЖАЙНОСТЬ, Ц/ГА-Очень большое: 5/5-{170.0000000, 409.0000000}" в модели: 4 "INF1"

| Код | Наименование класса | № | Код класса | Наименование класса | Сходство |
|-----|---|-----|------------|---|----------|
| 1 | ТИП ЦВЕТКА (151)-1/2:3 | 1 | 12 | УРОЖАЙНОСТЬ, Ц/ГА-Очень большое: 5/5-{170.0000000, 409.0000000} | 100.000 |
| 2 | ТИП ЦВЕТКА (151)-2/2:4 | 2 | 98 | ПЕРИОД СОЗРЕВАНИЯ (629)-6/7:7 | 40.236 |
| 3 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Очень малое: 1/... | 3 | 81 | ЯГОДЫ ОКРАСКА (225)-5/6:6 | 38.459 |
| 4 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Малое: 2/5-(112... | 4 | 66 | ЯГОДЫ РАЗМЕР (220)-3/5:6 | 37.895 |
| 5 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Среднее: 3/5-(17... | 5 | 2 | ТИП ЦВЕТКА (151)-2/2:4 | 36.689 |
| 6 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Большое: 4/5-(22... | 6 | 7 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Очень большое: 5/5-{280.0000000, ...} | 36.289 |
| 7 | СРЕДНЯЯ МАССА ГРОЗДИ, Г-Очень большое: ... | 7 | 44 | ВЫРЕЗКИ НИЖНИЕ ЛИСТЬЕВ (094-2)-5/5:9 | 34.985 |
| 8 | УРОЖАЙНОСТЬ, Ц/ГА-Очень малое: 1/5-(9.000... | 8 | 43 | ВЫРЕЗКИ НИЖНИЕ ЛИСТЬЕВ (094-2)-4/5:7 | 29.462 |
| 9 | УРОЖАЙНОСТЬ, Ц/ГА-Малое: 2/5-(60.0000000, ... | 9 | 20 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Среднее: 3/5-{6.4000000, 6.90... | 28.466 |
| 10 | УРОЖАЙНОСТЬ, Ц/ГА-Среднее: 3/5-(95.00000... | 10 | 91 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603)-2/3:3 | 27.701 |
| 11 | УРОЖАЙНОСТЬ, Ц/ГА-Большое: 4/5-(120.0000... | 11 | 58 | ПЛОТНОСТЬ ГРОЗДИ (204)-1/6:1 | 27.261 |
| 12 | УРОЖАЙНОСТЬ, Ц/ГА-Очень большое: 5/5-(17... | 12 | 59 | ПЛОТНОСТЬ ГРОЗДИ (204)-2/6:2 | 27.009 |
| 13 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Очень малое: ... | 98 | 26 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-4/6:6 | -23.337 |
| 14 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Малое: 2/5-(1... | 99 | 87 | ЯГОДЫ АРОМАТ (236)-5/7:7 | -23.461 |
| 15 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Среднее: 3/5-... | 100 | 108 | РОСТ ПОБЕГОВ (351)-4/4:9 | -24.865 |
| 16 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Большое: 4/5-... | 101 | 80 | ЯГОДЫ ОКРАСКА (225)-4/6:5 | -25.073 |
| 17 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Очень большо... | 102 | 1 | ТИП ЦВЕТКА (151)-1/2:3 | -25.950 |
| 18 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Очень мало... | 103 | 16 | САХАРИСТОСТЬ, Г/100 КУБ. СМ-Большое: 4/5-{18.7000000, 20.5... | -26.254 |
| 19 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Малое: 2/5-... | 104 | 10 | УРОЖАЙНОСТЬ, Ц/ГА-Среднее: 3/5-(95.0000000, 120.0000000) ... | -27.121 |
| 20 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Среднее: 3/... | 105 | 96 | ПЕРИОД СОЗРЕВАНИЯ (629)-4/7:5 | -32.969 |
| 21 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Большое: 4/... | 106 | 92 | НАПРАВЛЕНИЕ ИСПОЛЬЗОВАНИЯ (603)-3/3:4 | -33.661 |
| 22 | КИСЛОТНОСТЬ ТИТРУЧЕМАЯ, Г/Л-Очень боль... | 107 | 9 | УРОЖАЙНОСТЬ, Ц/ГА-Малое: 2/5-(60.0000000, 95.0000000) ... | -34.969 |
| 23 | РАССЕЧЕННОСТЬ ЛИСТЬЕВ (072)-1/6:1 | 108 | 37 | ВЫРЕЗКИ ВЕРХНИЕ ЛИСТЬЕВ (094-1)-3/5:5 | -34.974 |

Помощь Abs Prc1 Prc2 Inf1 Inf2 Inf3 Inf4 Inf5 Inf6 Inf7 График ВКЛ.фильтр по клщкале ВЫКЛ.фильтр по клщкале Вписать в окно Показать ВСЕ

Рисунок 19. Табличная форма с результатом кластерного анализа классов

Имеется возможность сразу получить графическую диаграмму, соответствующую приведенной таблице. Однако в этом случае эта диаграмма будет нечитабельна из-за большого количества графических элементов. Поэтому перед отображением графической диаграммы кликнем на кнопке: «Вписать в окно» на появившейся экранной форме, приведенной на рисунке 19, зададим 12 отображаемых классов, нажмем ОК, а уже затем кликнем по кнопке: «График» (рисунок 20):

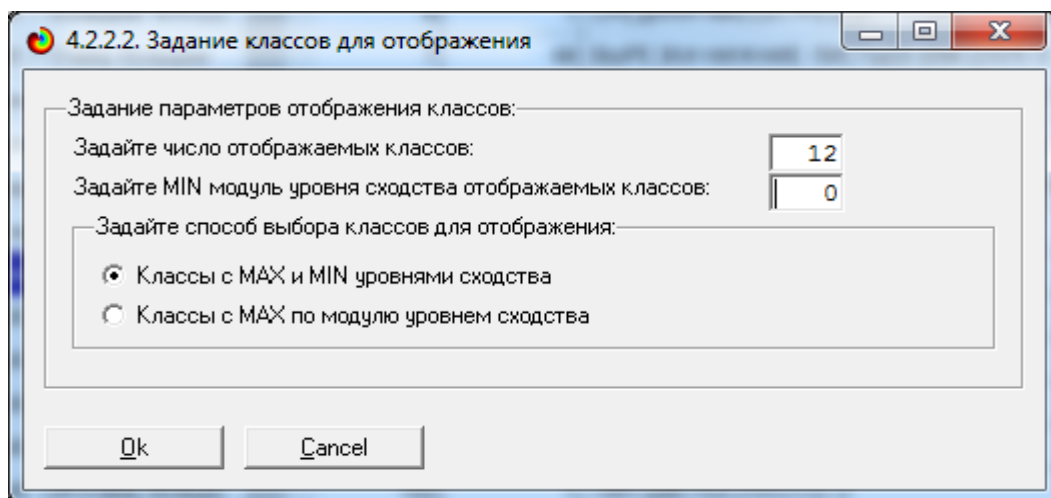


Рисунок 20. Задание параметров графической диаграммы с результатами кластерно-конструктивного анализа

А затем выведем саму диаграмму:

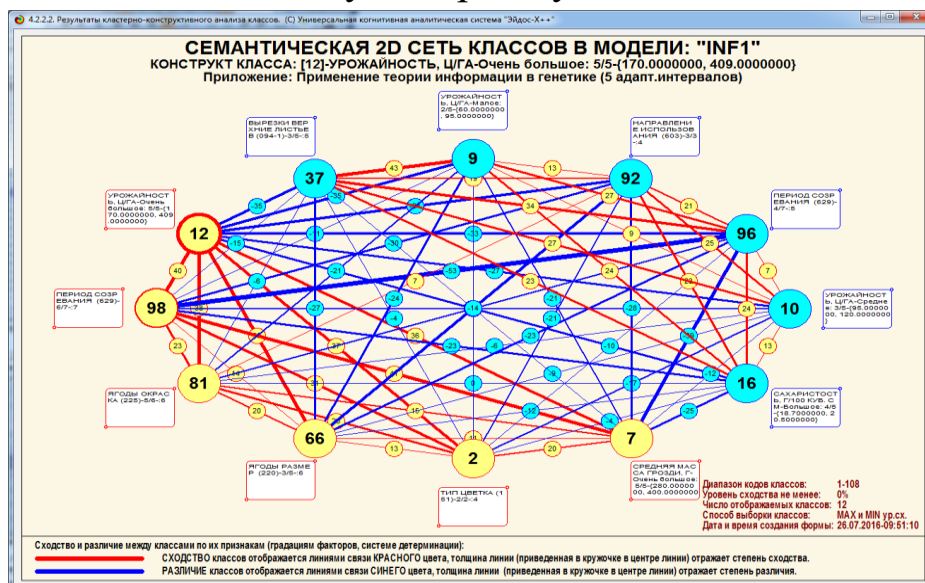


Рисунок 21. Когнитивная диаграмма с результатами кластерно-конструктивного анализа

На этой диаграмме изображен конструкт с двумя кластерами на полюсах. На одном полюсе кластер с центром в классе - код: 12 «Урожайность очень большая» (этот центр мы задаем в экранной форме, приведенной на рисунке 19) и классами 98, 81, 66 и 2, а на другом полюсе мы видим кластер с фенотипическими признаками: 37, 9, 92, 93, 10, 16.

Классы, находящиеся в одном кластере, детерминируются сходными системами генов и достижимы одновременно в одном сорте.

Классы, относящиеся к противоположным кластерам, находящимся на полюсах конструкта, не достижимы одновременно в одном сорте.

Например, перед селекционерами корректно ставить задачу получить высокоурожайный сорт винограда с окраской ягод 225 (код: 81) и размером ягод 220 (код: 66), и очень массивной гроздью (код: 7), при этом у этого сорта будет длительный период созревания 629 (код: 98). А вот требовать от селекционеров, чтобы эти ягоды были сладкими (код: 16) некорректно, т.к. это требует совершенно другого генома, чем обуславливающий ранее перечисленные фенотипические свойства. Если же мы хотим получить сладкий сорт, то, по-видимому, придется смириться с его малой или в лучшем случае средней урожайностью.

Отметим, что обычно подобными знаниями располагают только эксперты в данной предметной области, а в данной работе показано, как они могут быть получены с помощью современных автоматизированных интеллектуальных технологий (АСК-анализа и системы «Эйдос») непосредственно на основе эмпирических данных.

4.8. Количественное определение сходства-различия различных генов по их влиянию на фенотипические свойства. Возможность замены одних генов другими

Традиционно считается, что определенные гены обуславливают соответствующие фенотипические свойства и эти свойства не могут быть детерминированы другими генами. Но наше исследование показывает, что, по-видимому, это не так. Об этом говорит факт сходства между генами по их влиянию на фенотип, из-за чего они объединяются в кластеры, а противоположные кластеры образуют полюса конструкторов. В одном кластере оказываются гены, оказывающие сходное влияние на фенотипические свойства, а на полюсах конструктора – противоположное.

Для получения соответствующих выходных форм сначала сформируем матрицу сходства генов по их влиянию, а потом отобразим ее в табличном и графическом виде (рисунки 22 и 23):

| Код | Наименование признака | № | Код призна... | Наименование признака | Сходство |
|-----|---|-----|---------------|---|----------|
| 53 | VRZAG79-2-Среднее: 3/5-(249.0000000, 251.000... | 1 | 60 | VRZAG79-1-Очень большое: 5/5-(259.0000000, 261.0000000) | 100.000 |
| 54 | VRZAG79-2-Большое: 4/5-(251.0000000, 251.000... | 2 | 5 | VRZAG47-2-Очень большое: 5/5-(186.0000000, 195.0000000) | 43.409 |
| 55 | VRZAG79-2-Очень большое: 5/5-(251.0000000, 2... | 3 | 125 | VWMD27-2-Очень большое: 5/5-(186.0000000, 195.0000000) | 43.409 |
| 56 | VRZAG79-1-Очень малое: 1/5-(237.0000000, 251... | 4 | 167 | VRZAG67-1-Малое: 2/5-(135.0000000, 143.0000000) | 37.650 |
| 57 | VRZAG79-1-Малое: 2/5-(251.0000000, 251.00000... | 5 | 210 | WVIV67-1-Очень большое: 5/5-(376.0000000, 400.0000000) | 36.439 |
| 58 | VRZAG79-1-Среднее: 3/5-(251.0000000, 257.000... | 6 | 203 | WVIV67-2-Среднее: 3/5-(357.0000000, 358.0000000) | 32.063 |
| 59 | VRZAG79-1-Большое: 4/5-(257.0000000, 259.000... | 7 | 224 | VMC4F3.1-2-Большое: 4/5-(173.0000000, 177.0000000) | 28.423 |
| 60 | VRZAG79-1-Очень большое: 5/5-(259.0000000, 2... | 8 | 73 | VWMD32-2-Среднее: 3/5-(250.0000000, 252.0000000) | 26.310 |
| 61 | VWMD28-2-Очень малое: 1/5-(218.0000000, 234... | 9 | 173 | WVIV37-2-Среднее: 3/5-(156.0000000, 160.0000000) | 25.629 |
| 62 | VWMD28-2-Малое: 2/5-(234.0000000, 236.00000... | 10 | 83 | VWMD25-2-Среднее: 3/5-(239.0000000, 241.0000000) | 25.387 |
| 63 | VWMD28-2-Среднее: 3/5-(236.0000000, 236.0000... | 11 | 247 | WVIP31-1-Малое: 2/5-(181.0000000, 183.0000000) | 23.995 |
| 64 | VWMD28-2-Большое: 4/5-(236.0000000, 244.0000... | 12 | 16 | WV52-1-Очень малое: 1/5-(135.0000000, 141.0000000) | 22.362 |
| 65 | VWMD28-2-Очень большое: 5/5-(244.0000000, 25... | 240 | 85 | VWMD25-2-Очень большое: 5/5-(245.0000000, 255.0000000) | -19.451 |
| 66 | VWMD28-1-Очень малое: 1/5-(234.0000000, 244... | 241 | 71 | VWMD32-2-Очень малое: 1/5-(236.0000000, 248.0000000) | -20.255 |
| 67 | VWMD28-1-Малое: 2/5-(244.0000000, 258.00000... | 242 | 25 | VWMD7-2-Очень большое: 5/5-(247.0000000, 249.0000000) | -20.412 |
| 68 | VWMD28-1-Среднее: 3/5-(258.0000000, 258.0000... | 243 | 207 | WVIV67-1-Малое: 2/5-(361.0000000, 365.0000000) | -20.620 |
| 69 | VWMD28-1-Большое: 4/5-(258.0000000, 258.0000... | 244 | 226 | VMC4F3.1-1-Очень малое: 1/5-(161.0000000, 171.0000000) | -20.717 |
| 70 | VWMD28-1-Очень большое: 5/5-(258.0000000, 27... | 245 | 221 | VMC4F3.1-2-Очень малое: 1/5-(161.0000000, 165.0000000) | -21.794 |
| 71 | VWMD32-2-Очень малое: 1/5-(236.0000000, 248... | 246 | 49 | VRZAG62-1-Большое: 4/5-(202.0000000, 204.0000000) | -21.876 |
| 72 | VWMD32-2-Малое: 2/5-(248.0000000, 250.00000... | 247 | 153 | WVIQ52-2-Среднее: 3/5-(76.0000000, 78.0000000) | -22.986 |
| 73 | VWMD32-2-Среднее: 3/5-(250.0000000, 252.0000... | 248 | 146 | VMC1B11-1-Очень малое: 1/5-(167.0000000, 179.0000000) | -23.353 |
| 74 | VWMD32-2-Большое: 4/5-(252.0000000, 256.0000... | 249 | 26 | VWMD7-1-Очень малое: 1/5-(235.0000000, 243.0000000) | -24.103 |
| 75 | VWMD32-2-Очень большое: 5/5-(256.0000000, 27... | 250 | 103 | WVIB01-2-Среднее: 3/5-(291.0000000, 295.0000000) | -26.525 |

Рисунок 22. Результат кластерно-конструктивного анализа генов по их влиянию на фенотип

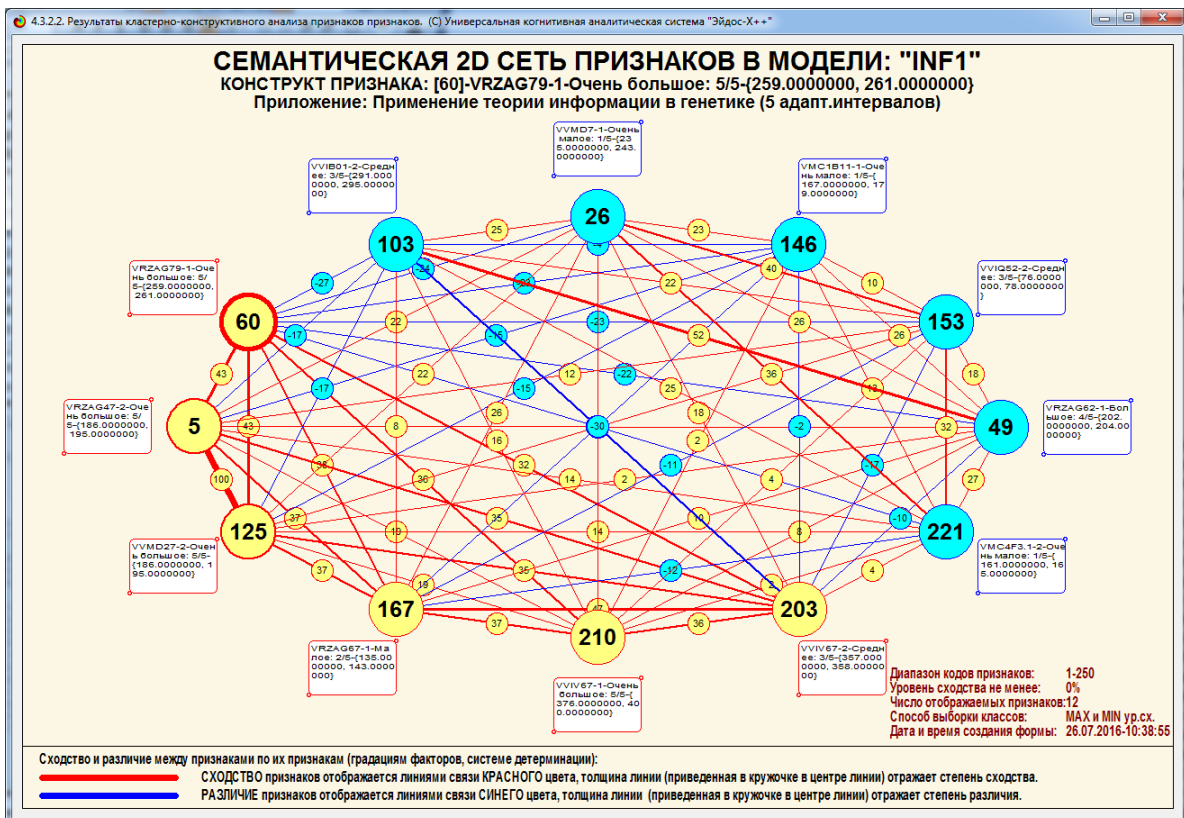


Рисунок 23. Результат кластерно-конструктивного анализа генов по их влиянию на фенотип

Из этих выходных форм видно, какие гены оказывают сходное, а какие - противоположное влияние на фенотип. Если бы между генами и фенотипическими признаками/свойствами было взаимно-однозначное соответствие, то между ними не было бы сходства и различия по их влиянию на фенотип, т.е. их система была бы ортонормированна.

Отметим, что обычно подобными знаниями располагают только эксперты в данной предметной области, а в данной работе показано, как они могут быть получены с помощью современных автоматизированных интеллектуальных технологий (АСК-анализа и системы «Эйдос») непосредственно на основе эмпирических данных.

4.9. Алгоритм принятия решения о структуре генома с целью создания нового сорта с заданной системой фенотипических свойств/признаков

Авторы предлагают следующий алгоритм принятия решения о структуре генома с целью создания нового сорта с заданной системой фенотипических свойств/признаков, по сути совпадающий с логикой изложения в данной работе (предполагающий, что все этапы АСК-анализа до создания и верификации моделей успешно выполнены):

Шаг-1: 4.7. Количественное определение сходства-различия различных фенотипических свойств по их системе детерминации генами. Выбираем совместимые в одном сорте фенотипические свойства.

Шаг-2: 4.5. Провести автоматизированный SWOT-анализ системы детерминации каждого из выбранных на 1-м шаге фенотипического свойства структурой генома. Выбрать гены, детерминирующие выбранные фенотипические свойства.

Шаг-3: Исключить повторы генов, выбранных на 2-м шаге.

Шаг-4: 4.4. Решить задачу системной идентификации сорта по выбранным на 3-м шаге генам. Этот сорт и берем за основу для селекции.

Для выполнения 4-го шага создаем еще одну модель совершенно аналогично описанному выше, но в качестве классов выбираем наименования сортов, а в качестве признаков – гены. Авторами эта модель создана. На рисунке 24 приводится одна из форм, отражающая результаты кластерно-конструктивного анализа сравнения сортов по их геному в этой модели:

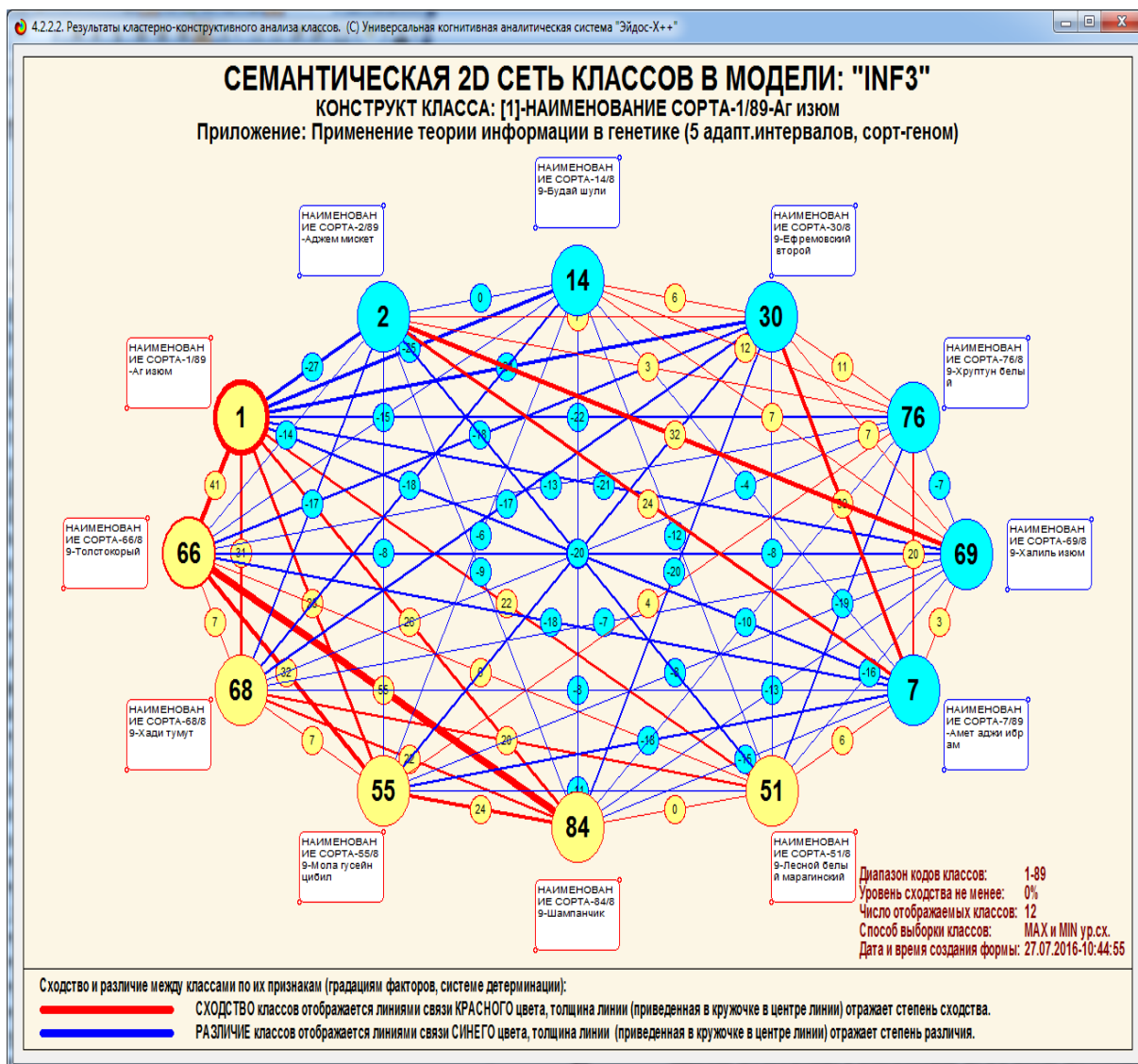


Рисунок 24. Результаты кластерно-конструктивного анализа сравнения сортов по их геному

4.10. Когнитивные функциональные зависимости между структурой генома и степенью выраженности фенотипических свойств/признаков

Когнитивные функции – это предложенное проф. Е.В. Луценко обобщение классического понятия функциональной зависимости на основе теории информации в системной нечеткой интервальной математике [20]. Для их генерации и визуализации в системе «Эйдос» предназначен режим 4.5 (рисунок 25):



Рисунок 25. Начальная экранная форма режима визуализации когнитивных функций

Модуль визуализации когнитивных функций разработан по постановке проф. Е.В. Луценко разработчиком интеллектуальных систем из Белоруссии Д.К. Бандык [21].

На рисунке 26 приведена визуализации всей базы знаний INF1 [8], а на рисунке 28 – примеры некоторых когнитивных функций:

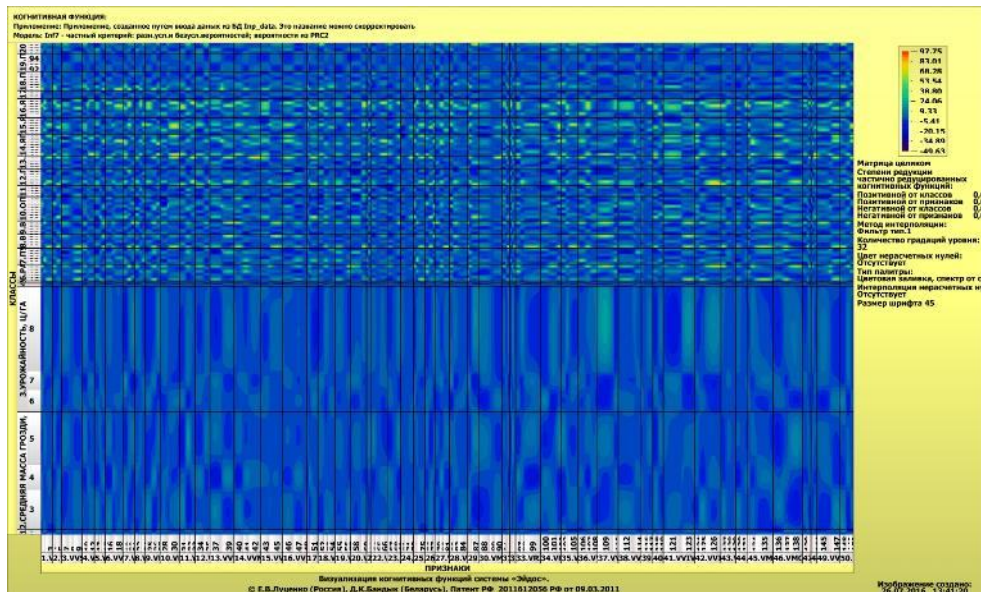


Рисунок 26. Когнитивная функция всех подматриц модели INF1

Всего в каждой из 7 моделей знаний получается подобных 1000 когнитивных функций, т.к. в каждой модели знаний 20 классификационных шкал и 50 описательных шкал, т.е. 1000 подматриц (на рисунке они отделены черными линиями).

Если бы в предметной области не было ярко выраженных закономерностей, то изображение на рисунке напоминало бы изображение на экране телевизора при отсутствии сигнала (случайная рябь). На рисунке приведено изображение базы знаний модели INF1 случайной модели [22]¹⁶, созданной на основе модели, представленной на рисунке 26.

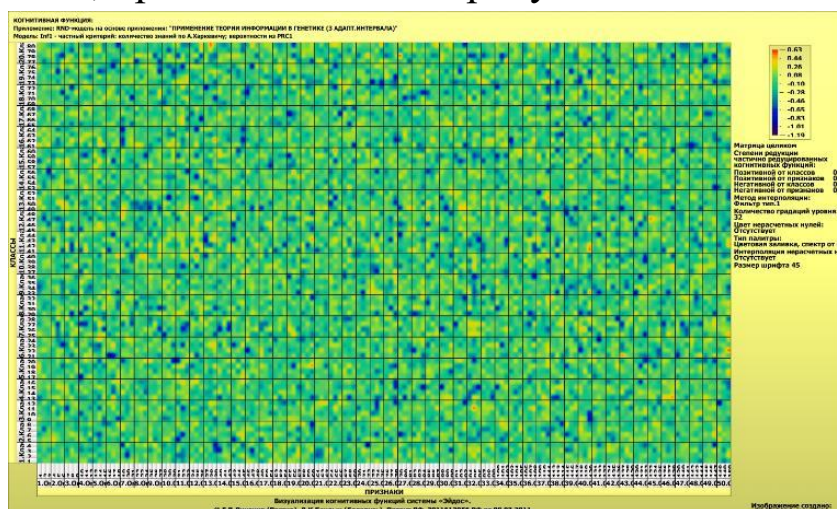
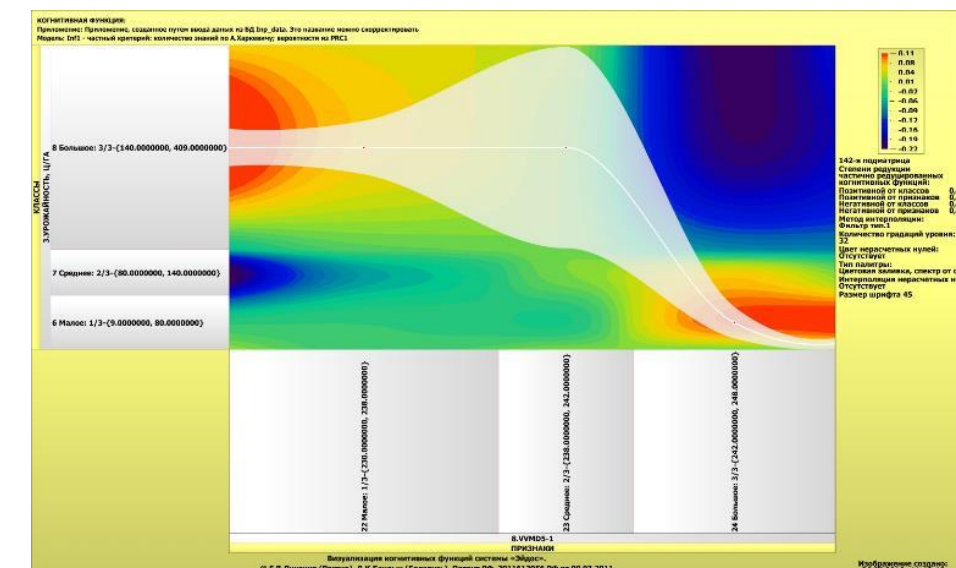
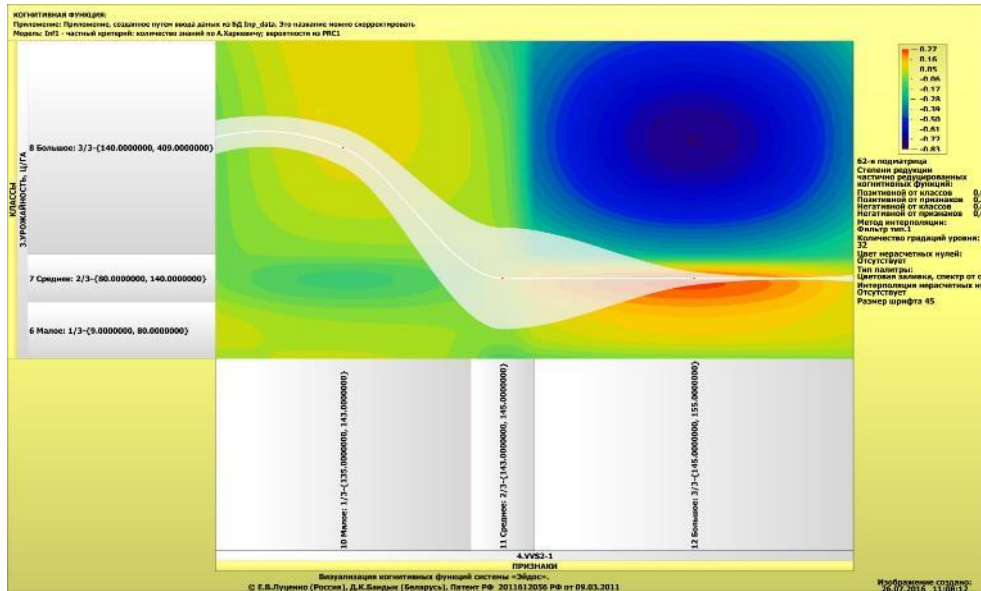
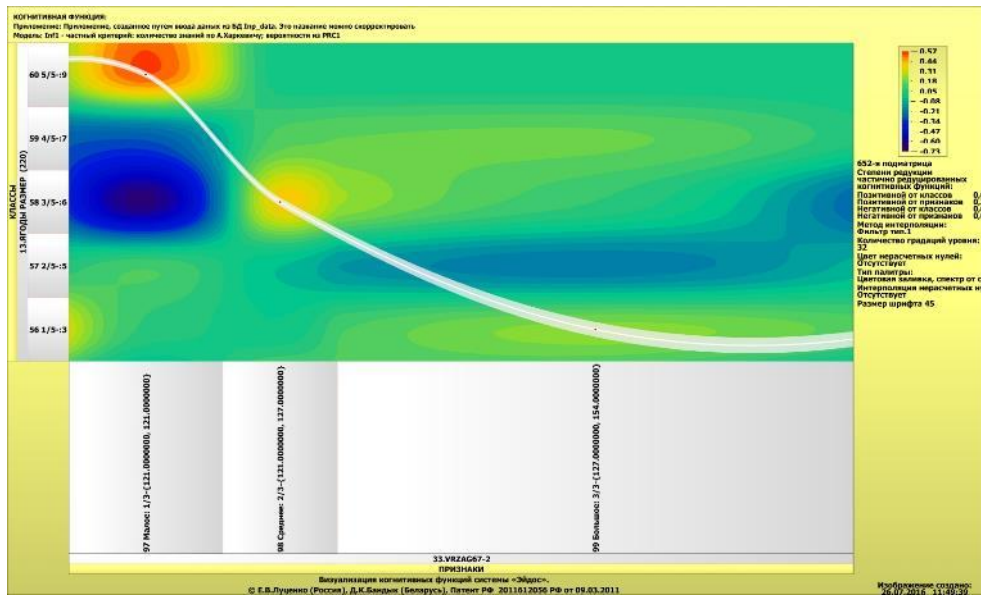
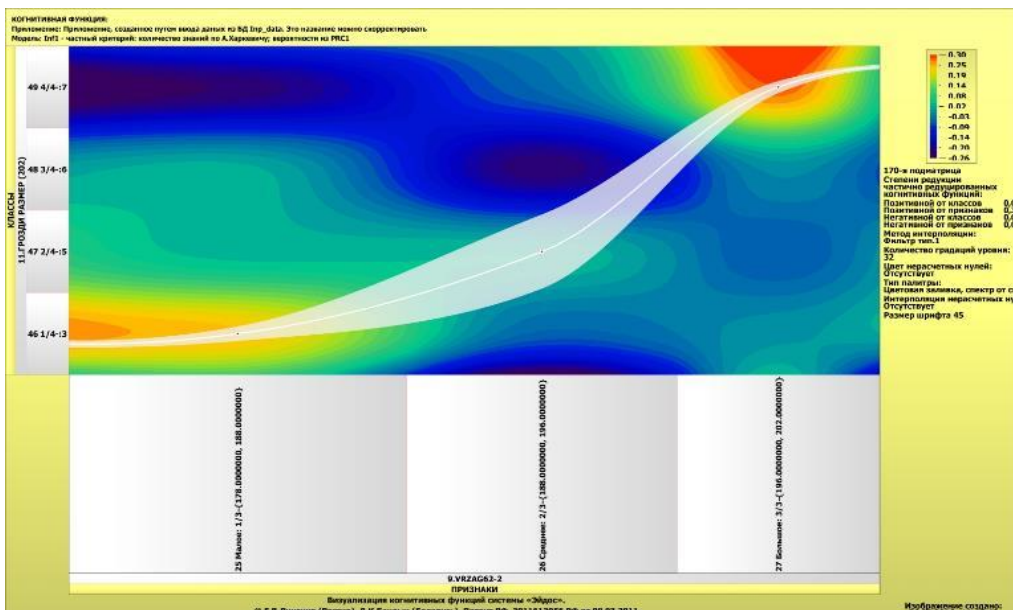
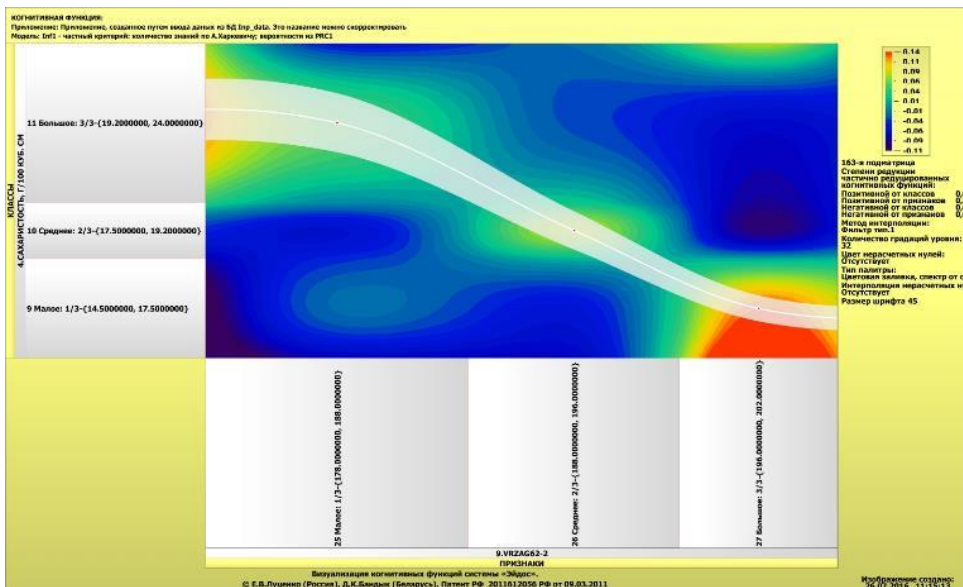
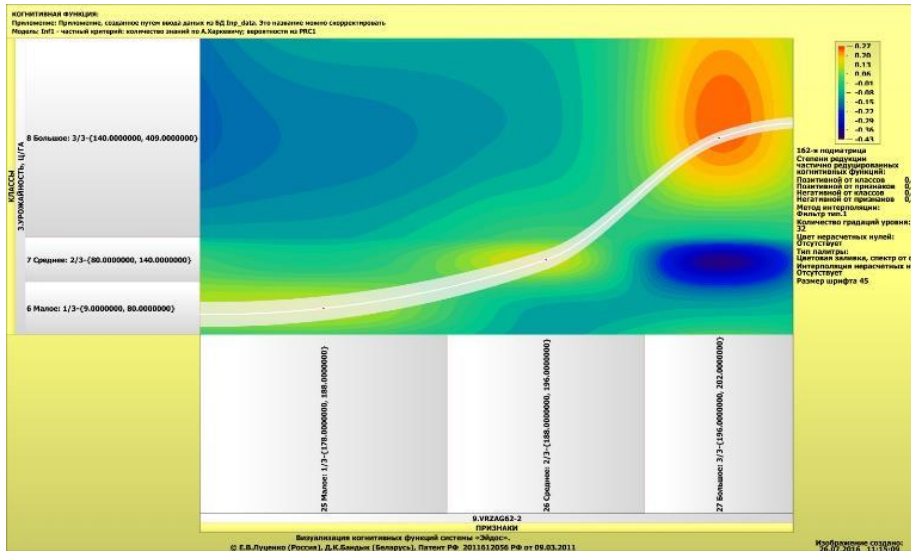


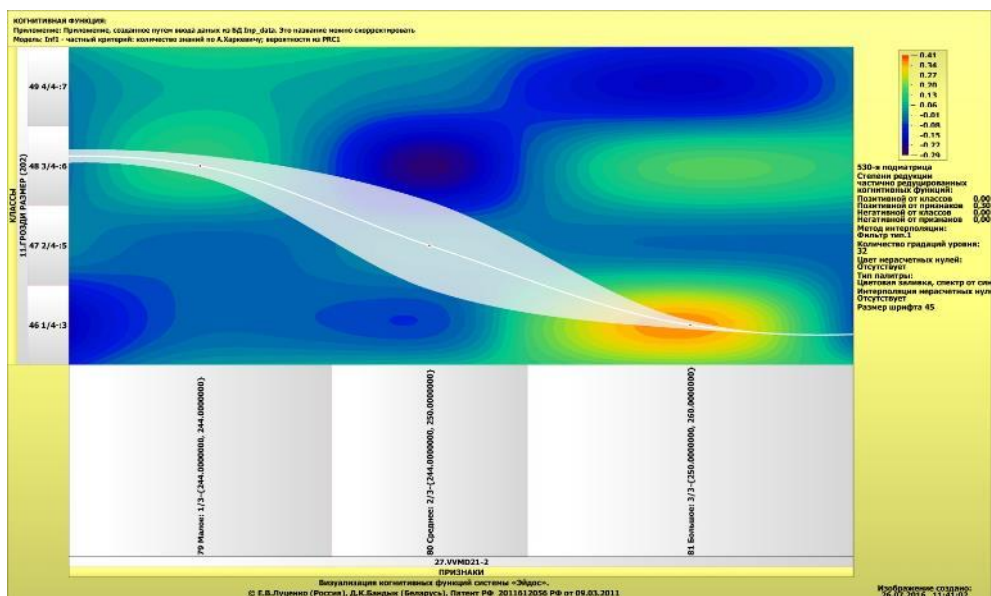
Рисунок 27. Когнитивная функция всех подматриц случайной модели INF1

¹⁶ См., например: http://lc.kubagro.ru/aidos/aidos06_lab/lab_10.htm

Различие рисунков 26 и 27 очевидно.







Рисунки 28. Примеры некоторых когнитивных функций

Эти когнитивные функции отражают зависимости степени выраженности различных фенотипических свойств/признаков винограда от степени выраженности определенных генов.

5. Выводы и перспективы

Общеизвестно, что генетика изучает механизмы наследственности/изменчивости и очень широко пользуется понятием «наследственная информация». При этом генетика под информацией подразумевает содержание генетического кода - структуры молекул ДНК и РНК, входящих в состав клетки живого организма. Генетика изучает механизмы записи, копирования, считывания генетической информации, возможности ее модификации, а также ее влияние на свойства/признаки организма. В разговорном и научном языке прочно закрепились фразы, типа «Гены содержат информацию о признаках и свойствах организма». Парадоксально, но мы не видим попыток определения количества информации, содержащейся в конкретных генах о конкретных фенотипических свойствах/признаках организма. Казалось бы, применение теории

информации в генетике является совершенно естественным и напрашивается сама собой. Тем более странно, что практически нет работ, посвященных применению теории информации для решения задач генетики. Данная статья призвана в какой-то степени восполнить этот пробел на примере вычисления количества информации в генах о свойствах/признаках различных сортов винограда. В ней рассматривается применение автоматизированного системно-когнитивного анализа (АСК-анализ), его математической модели – системной теории информации и реализующего их программного инструментария – интеллектуальной системы «Эйдос» для решения одной из важных задач генетики: определения количества информации, содержащейся в генах о различных фенотипических свойствах/признаках винограда. Для решения этой задачи выполняются следующие этапы: 1) когнитивно-целевая структуризация предметной области; 2) формализация предметной области, т.е. разработка классификационных и описательных шкал и градаций и обучающей выборки; 3) синтез и верификация информационной модели, отражающей количество информации в генах о фенотипических свойствах/признаках (многопараметрическая типизация); 4) вывод информации о генетической системе детерминации фенотипических свойств/признаков (SWOT-анализ феносвойств); 5) вывод информации о силе и направлении влияния конкретного гена на фенотипические свойства/признаки (SWOT-диаграммы генов); 6) решение задачи системной идентификации фенотипических свойств/признаков по наличию тех или иных генов; 7) количественное определение сходства-различия различных фенотипических свойств/признаков по их системе детерминации генами. Конкретное фенотипическое свойство/признак рассматривается как зашумленный генетический текст, включающий как генетическую информацию об истинном феносвойстве/признаке (чистый сигнал), так и шум,

искажающий эту информацию, обусловленный случайным воздействием окружающей среды. Программный инструментальный АСК-анализа – интеллектуальная система «Эйдос» - обеспечивает подавление шума и выделение истинного сигнала.

Описанная в статье технология синтеза и применения интеллектуальной измерительной системы может быть применена не только при генетических исследованиях в ампелографии, но и в других предметных областях, чему может способствовать и то, что система «Эйдос» размещена в полном открытом бесплатном доступе на сайте первого автора по адресу: <http://lc.kubagro.ru/aidos/Aidos-X.htm>.

Конечно, данная работа далека от завершенности и лишь продемонстрировала возможность применения теории информации и когнитивных технологий в исследованиях в области генетики. В перспективе авторы планируют провести более масштабное исследование на значительно большем объеме исходных данных: большем количестве сортов винограда, большем количестве их фенотипических свойств/признаков и большем количестве генов, причем не только резервных, микросателлитных, но и оказывающих непосредственное влияние на фенотипические свойства/признаки.

Материалы данной статьи могут быть использованы при проведении лабораторных работ по дисциплинам, связанным с интеллектуальными технологиями, представлением знаний и системами искусственного интеллекта, а также по эконометрике, биометрии, экологии, педагогике, психологии, медицине, криминалистике, энтомологии и др.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. 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. <http://elibrary.ru/item.asp?id=18632909>.

2. Орлов А.И., Луценко Е.В. Системная нечеткая интервальная математика. Монография (научное издание). – Краснодар: КубГАУ, 2014, – 600 с. ISBN 978-5-94672-757-0. <http://elibrary.ru/item.asp?id=21358220>.
3. Луценко Е.В. Универсальная когнитивная аналитическая система «Эйдос». Монография (научное издание). – Краснодар: КубГАУ, 2014. – 600 с. ISBN 978-5-94672-830-0. <http://elibrary.ru/item.asp?id=22401787>.
4. Луценко Е.В. Исследование влияния подсистем различных уровней иерархии на эмерджентные свойства системы в целом с применением АСК-анализа и интеллектуальной системы "Эйдос" (микроструктура системы как фактор управления ее макросвойствами) / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2012. – №01(075). С. 638 – 680. – Шифр Информрегистра: 0421200012\0025, IDA [article ID]: 0751201052. – Режим доступа: <http://ej.kubagro.ru/2012/01/pdf/52.pdf>, 2,688 у.п.л.
5. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. - М.: Высшая школа, 1989. - 320 с.
6. Перегудов Ф.И., Тарасенко Ф.П. Основы системного анализа. – Томск: Изд-во науч.-техн. лит., 1997. – 389 с.
7. Луценко Е.В. Теоретические основы, технология и инструментарий автоматизированного системно-когнитивного анализа и возможности его применения для сопоставимой оценки эффективности вузов / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №04(088). С. 340 – 359. – IDA [article ID]: 0881304022. – Режим доступа: <http://ej.kubagro.ru/2013/04/pdf/22.pdf>, 1,25 у.п.л.
8. Луценко Е.В. Синтез адаптивных интеллектуальных измерительных систем с применением АСК-анализа и системы «Эйдос» и системная идентификация в эконометрике, биометрии, экологии, педагогике, психологии и медицине / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1 – 60. – IDA [article ID]: 1161602001. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/01.pdf>, 3,75 у.п.л.
9. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системно-когнитивном анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №08(092). С. 859 – 883. – IDA [article ID]: 0921308058. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>, 1,562 у.п.л.
10. Стабин И.П., Моисеева В.С. Автоматизированный системный анализ. - М.: Машиностроение, 1984. - 309 с.
11. Симанков В.С. Автоматизация системных исследований в альтернативной энергетике. Диссерт. на соиск. уч. ст. докт, техн. наук. По спец.: 05.13.01. <http://tekhnosfera.com/avtomatizatsiya-sistemnyh-issledovaniy-v-alternativnoy-energetike>.
12. Klir G.J. Architecture of Systems Problem Solving, with D. Elias. – New York.: Plenum Press, 1974. –354 p.
13. Клир Дж. Системология. Автоматизация решения системных задач. - Москва: Радио и связь, 1990. - 538 с. <http://www.twirpx.com/file/486296/>.
14. Лефевр В.А. Конфликтующие структуры . Издание второе, переработанное и дополненное. — М.: Изд-во «Советское радио», 1973. – 158 с. с ил.

15. Хаббард Дуглас У. Как измерить все, что угодно. Оценка стоимости нематериального в бизнесе / Дуглас У. Хаббард / [Пер. с англ. Е. Пестеревой]. — М.: ЗАО «Олимп-Бизнес», 2009. — 320 с.: ил. ISBN 978-5-9693-0163-4 (рус.). <http://www.twirpx.com/file/1546361/>.
16. Сайт автора АСК-анализа проф. Е.В. Луценко: <http://lc.kubagro.ru/>.
17. Трошин Л.П., Маградзе Д.Н. Ампелографический скрининг генофонда винограда (Учебное наглядное пособие). – Краснодар, 2013. – 120 с.
18. Луценко Е.В. Количественный автоматизированный SWOT- и PEST-анализ средствами АСК-анализа и интеллектуальной системы «Эйдос-X++» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №07(101). С. 1367 – 1409. – IDA [article ID]: 1011407090. – Режим доступа: <http://ej.kubagro.ru/2014/07/pdf/90.pdf>, 2,688 у.п.л.
19. Луценко Е.В. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе «Эйдос») / Е.В. Луценко, В.Е. Коржаков // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2011. – №07(071). С. 528 – 576. – Шифр Информрегистра: 0421100012\0253, IDA [article ID]: 0711107040. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>, 3,062 у.п.л.
20. Луценко Е.В. Когнитивные функции как обобщение классического понятия функциональной зависимости на основе теории информации в системной нечеткой интервальной математике / Е.В. Луценко, А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №01(095). С. 122 – 183. – IDA [article ID]: 0951401007. – Режим доступа: <http://ej.kubagro.ru/2014/01/pdf/07.pdf>, 3,875 у.п.л.
21. Луценко Е.В., Бандык Д.К. Подсистема визуализации когнитивных (каузальных) функций системы «Эйдос» (Подсистема «Эйдос-VCF»), Пат. № 2011612056 РФ. Заяв. № 2011610347 РФ 20.01.2011. Опубл. от 09.03.2011. <http://lc.kubagro.ru/aidos/2011612056.jpg>
22. Луценко Е.В. Лабораторный практикум по интеллектуальным информационным системам: Учебное пособие для студентов специальности "Прикладная информатика (по областям)" и другим экономическим специальностям. 2-е изд., перераб. и доп. – Краснодар: КубГАУ, 2006. – 318 с. <http://elibrary.ru/item.asp?id=21683721>
23. Биометрическая оценка полиморфизма сортогрупп винограда Пино и Рислинг по морфологическим признакам листьев среднего яруса кроны / Л.П. Трошин, Е.В. Луценко, П.П. Подваленко, А.С. Звягин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №08(052). С. 1 – 14. – Шифр Информрегистра: 0420900012\0097, IDA [article ID]: 0520908001. – Режим доступа: <http://ej.kubagro.ru/2009/08/pdf/01.pdf>, 0,875 у.п.л.
24. Луценко Е.В. Решение задач ампелографии с применением АСК-анализа изображений листьев по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко, Д.К. Бандык, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №08(112). С. 862 – 910. – IDA [article ID]: 1121508064. – Режим доступа: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>, 3,062 у.п.л.

25. Луценко Е.В. Количественное измерение сходства-различия клонов винограда по контурам листьев с применением АСК-анализа и системы «Эйдос» / Е.В. Луценко, Л.П. Трошин, Д.К. Бандык // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №02(116). С. 1205 – 1228. – IDA [article ID]: 1161602077. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/77.pdf>, 1,5 у.п.л.
26. Анапская ампелографическая коллекция / Е.А. Егоров, О.М. Ильяшенко, А.Г. Коваленко [и др.]. – Краснодар: СКЗНИИСиВ, 2009. – 216 с.
27. Грамотенко П.М., Трошин Л.П. Микросистематика винограда: классификация сортов винограда А.М. Негруля и ее дальнейшее развитие // Виноградарство и виноделие. – 1994. – № 1. – С. 10–17.
28. Интерактивная ампелография и селекция винограда: (Сб. материалов Междунар. симп.). – Краснодар: КубГАУ, 2012. – 264 с.
29. Операционные данные ампелографической коллекции "Магарац" за 1985–1987 гг. / Л.П. Трошин, А.М. Панарина, А.М. Пискарева [и др.]; ВНИИВиПП "Магарац". – Ялта, 1988. – 86 с.
30. Трошин Л.П. Ампелография и селекция винограда. – Краснодар: РИЦ «Вольные мастера», 1999. – 138 с.
31. Трошин Л.П. Лучшие сорта винограда Евразии. – Краснодар: Алви-Дизайн, 2006. – 224 с.
32. Трошин Л. П., Федоров Ю. К. Биометрический анализ генофонда винограда / ВНИИВиПП "Магарац". – Ялта, 1988. – 90 с.
33. Трошин Л.П. Ампелографическая и селекционная научно-исследовательская работа Кубанского госагроуниверситета [Электронный ресурс] // Науч. журн. КубГАУ. – Краснодар: КубГАУ, 2012. – № 07 (81). – С. 524–544. – Режим доступа: <http://ej.kubagro.ru/2012/07/pdf/39.pdf>, 1,312 у.п.л.
34. Трошин Л.П., Радчевский П.П. Виноград: иллюстрированный каталог. Районированные, перспективные, тиражные сорта. – Ростов н/Д: Феникс, 2010. – 271 с.: ил. – (Мир садовода).
35. Website: <http://www.bioversityinternational.org/fileadmin/bioversity/publications/pdfs/393.pdf>.
36. Website: <http://www.diprove.unimi.it/GRAPENET/index.php>.
37. Website: <http://www.eu-vitis.de/index.php> и <http://www.vivc.de/index.php>.
38. Website: <http://www.oiv.int/oiv/info/frplublicationoiv#desc>.
39. Website <http://plantgrape.plantnet-project.org/> и <http://lescepages.free.fr/cepmc.html>.
40. Website: <http://www.vitis.ru> и <http://kubsau.ru/chairs/viniculture/>.
41. Изучение генетического разнообразия генофона винограда Северного Кавказа / Р. Тёпфер, Э. Мауль, А.В. Милованов и др. // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №05(119). С. 1337 – 1355. – IDA [article ID]: 1191605090. – Режим доступа: <http://ej.kubagro.ru/2016/05/pdf/90.pdf>, 1,188 у.п.л.
42. Милованов А.В. Генотипирование новых перспективных технических протоклонов винограда с использованием микросателлитных маркёров / А.В. Милованов, А.С. Звягин, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. –

№04(098). С. 139 – 157. – IDA [article ID]: 0981404010. – Режим доступа: <http://ej.kubagro.ru/2014/04/pdf/10.pdf>, 1,188 у.п.л., импакт-фактор РИНЦ=0,346.

43. Милованов А.В. Генотипирование сортов винограда по молекулярным маркерам / А.В. Милованов, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №02(096). С. 53 – 65. – IDA [article ID]: 0961402005. – Режим доступа: <http://ej.kubagro.ru/2014/02/pdf/05.pdf>, 0,812 у.п.л., импакт-фактор РИНЦ=0,346.

44. Трошин Л.П., Милованов А.В., Звягин А.С. Ампело-генетический скрининг перспективных столовых и технических сортов и протоклонов винограда // Виноделие и виноградарство. – 2015. - № 2. – С. 33-36.

45. Трошин Л.П., Милованов А.В., Звягин А.С. Этюд совершенствования клоновой селекции винограда // Магарач. Виноградарство и виноделие. – 2015. - №3. – С. 33-36.

46. Хильчевская Р.И. Роль асимметрии-симметрии материи в процессах происхождения жизни на Земле // Ж-л ВХО им. Д.И. Менделеева, Том XXV4, 1980 г. - Изд-во. «Химия». - С. 418–424. Адрес доступа: <http://omdp.narod.ru/gip/rolasim.htm>

47. Ленинджер А.. В кн.: Биохимия. - М., Мир, 1974. - С. 869, 655, 661.



Генет ПРЕОБРАЖЕНИЕ



Генет ХРИЗОЛИТ

5. БИОМЕТРИЧЕСКАЯ ОЦЕНКА ПОЛИМОРФИЗМА СОРТОГРУПП ВИНОГРАДА ПИНО И РИСЛИНГ ПО МОРФОЛОГИЧЕСКИМ ПРИЗНАКАМ ЛИСТЬЕВ СРЕДНЕГО ЯРУСА КРОНЫ*

В ампелографии полиморфизмом называют биологическое явление разнообразия - одновременное перманентное наличие в популяции двух или нескольких фенотипически и генотипически отличающихся групп растений одного сорта, то есть клонов. В виноградарстве обычно считают клон изогенной популяцией, а сорт как совокупность клонов или форм – гетерогенной. Полиморфизм сортов имеет большое практическое значение, так как сорта-клоны обеспечивают при условии их гармоничного взаимодополняющего сочетания адаптивный ампеоценоз и высокий для производства экономический эффект. В статье на фенотипическом уровне по листьям освещен полиморфизм двух гетерогенных популяций – сортогрупп Пино и Рислинг, являющихся лучшими представителями французского и немецкого виноградарства.

Введение

История развития виноградарства показывает, что большинство возделываемых сортов винограда возникло в результате того, что человек отбирал и размножал отклонившиеся по каким-то признакам формы, вводя их в культуру. На протяжении многих веков сорта наполнялись различными рекомбинантами, мутантами и клонами, образуя соответствующие сортогруппы, сортотипы и эколого-географические группы [1, 5, 10].

Сложившийся в настоящее время основной российский сортимент возделываемых сортов, в основном, представлен сортами-интродуцентами, завезенными из-за рубежа. Как показали последние исследования ампелографов, они оказались полиморфными и гетерогенными, что вызывает нестабильность урожайности и сильное варьирование качества урожая [8]. Примеры полиморфизма: наличие сортов с функционально

женским типом цветка среди одноименных обоеполюх, форм с петрушечными листьями у Шаслы белой, появление тетраплоидов у диплоидного Рислинга, Кокура красного среди Кокура белого, Муската розового среди Муската белого и другие мутационные изменения фенотипов, сохраняющихся в вегетативных поколениях.

Материал и методы

Объектами исследований являлись гетерогенные популяции двух сортогрупп, в ампелографической литературе называемых Пино и Рислинг, всего 231 учетный куст. Из них известный французский сорт Пино белый (контроль) насчитывал 35 кустов, его клоны Пинок белый, Пино белый № 32 и Пино черный № 07 – от 12 до 56; другой сорт немецкого происхождения Рислинг (рейнский, контроль) имел 15 кустов, его клоны Рислиналк, Рислинг Алькадар № 34А, Рислинг Алькадар № 34Б, Рислинг № 4-9-2 и Рислинг анапский – от 14 до 34 кустов. Все перечисленные клоны являлись вегетативным потомством высокопродуктивных протоклонов, которые в свое время были отобраны первым автором статьи в корнесобственных старых насаждениях сортов Пино и Рислинг, некогда произраставших в двух местах - Бахчисарайском районе и под городом Севастополем Республики Крым.

Исследуемые генотипы (сорта и клоны) произрастали на одном опытном участке в учхозе «Кубань» КубГАУ, были одновозрастными и с ними проводились одинаковые уходные работы. Исследования проводились в течение трех лет, с 2006 по 2008 гг., по данным измерений 10 листьев среднего яруса кроны каждого куста.

Культура возделывания этих генотипов укрывная, корнесобственная, кусты сформированы по типу бесштамбового многорукавного веера. Схема посадки корнесобственных

насаждений $3,0 \times 1,0$ м, год посадки – 1999. Участки не орошаемые. Технология возделывания является общепринятой для южной зоны промышленного виноградарства РФ.

Описывались и сравнивались листья по следующему комплексу из десяти признаков:

- длина черешка;
- длина листовой пластинки – от верхней точки зубца центральной лопасти до нижнего зубца нижней лопасти;
- ширина листовой пластинки – между выступающими зубцами средних жилок боковых лопастей;
- длина срединной жилки;
- расстояние от верхнего бокового выступа до черешковой выемки;
- расстояние от нижнего бокового выступа до черешковой выемки;
- добухтовое верхнее расстояние – от дна верхней боковой вырезки до черешковой выемки;
- добухтовое нижнее расстояние – от дна нижней боковой вырезки до черешковой выемки;
- угол α – угол между центральной жилкой и выступающей верхней боковой лопастью;
- угол β – угол между центральной жилкой и выступающей нижней боковой лопастью.

Математико-статистическую обработку экспериментальных данных проводили по методикам, описанным Е.В. Луценко [2], В.О. Островерховым и Л.П. Трошиным [3], Ю.Н. Тюриным и А.А. Макаровым [9].

Результаты исследований

Очень важным вопросом селекционно-генетических исследований виноградной лозы является изучение полиморфизма сортогрупп по морфологическим признакам

листа. Связано это с проблемой использования морфометрии листьев как выражения фенотипического полиморфизма винограда. Полиморфные группы растений как наследственно новые вариации могут отличаться от исходной популяции по большому числу количественных признаков, определяемых интегрированной системой генов. Не всегда отличия по каждому отдельно взятому признаку велики, однако суммарно они могут быть выделены на основе анализа комплекса признаков. Именно это и требуется при отборе высокопродуктивных клонов, т.к. необходимо улучшение сорта по комплексу признаков урожайности, устойчивости к неблагоприятным условиям, периоду вегетации, сахаронакоплению и др. При этом следует ориентироваться на поиск растений, отклоняющихся от средних популяционных значений признаков, так как среди них более вероятно нахождение мутантов и редких комбинаций генов [4-7].

Поэтому нами последовательно изучалась внутрисортная и межсортная изменчивость по усредненным за три года измерениям 10 признаков листа, включающим линейные и угловые параметры, которые были зафиксированы у каждого учетного куста (табл. 1).

Таблица 1. - Анализ морфологических признаков листьев исследуемых генотипов винограда в сортогруппе Пино

| Генотипы | *X | *S _x | *CV | *d | *P _b , % |
|-------------------------------|-------|-----------------|-------|------|---------------------|
| Длина листовой пластинки, см | | | | | |
| Пино белый | 15,87 | 0,15 | 9,20 | | |
| Пино черный № 07 | 15,66 | 0,15 | 7,79 | 0,21 | 83,7 |
| Пинок белый | 14,94 | 0,18 | 8,4 | 0,93 | 99,9 |
| Пино белый № 32 | 15,73 | 0,22 | 10,52 | 0,14 | 71,4 |
| Ширина листовой пластинки, см | | | | | |
| Пино белый | 15,99 | 0,15 | 8,78 | | |
| Пино черный № 07 | 15,94 | 0,17 | 8,67 | 0,05 | 59,2 |
| Пинок белый | 14,80 | 0,15 | 6,93 | 1,19 | 100,0 |
| Пино белый № 32 | 15,31 | 0,23 | 11,70 | 0,68 | 99,4 |
| Длина черешка, см | | | | | |
| Пино белый | 8,59 | 0,17 | 18,49 | | |

| | | | | | |
|--|--------|------|-------|-------|-------|
| Пино черный № 07 | 8,69 | 0,14 | 13,32 | 0,10 | 67,6 |
| Пинок белый | 7,51 | 0,17 | 15,42 | 1,08 | 100,0 |
| Пино белый № 32 | 7,93 | 0,24 | 23,04 | 0,66 | 99,0 |
| Длина срединной жилки, см | | | | | |
| Пино белый | 10,84 | 0,12 | 10,27 | | |
| Пино черный № 07 | 10,72 | 0,10 | 7,96 | 0,12 | 76,8 |
| Пинок белый | 10,42 | 0,13 | 8,51 | 0,42 | 99,3 |
| Пино белый № 32 | 10,36 | 0,14 | 10,10 | 0,48 | 99,6 |
| Расстояние от верхнего бокового выступа до черешковой выемки, см | | | | | |
| Пино белый | 10,05 | 0,11 | 10,72 | | |
| Пино черный № 07 | 9,76 | 0,11 | 8,96 | 0,29 | 97,0 |
| Пинок белый | 9,16 | 0,12 | 8,81 | 0,89 | 100,0 |
| Пино белый № 32 | 9,77 | 0,14 | 10,71 | 0,28 | 94,1 |
| Расстояние от нижнего бокового выступа до черешковой выемки, см | | | | | |
| Пино белый | 7,64 | 0,09 | 10,94 | | |
| Пино черный № 07 | 7,02 | 0,08 | 9,86 | 0,62 | 100,0 |
| Пинок белый | 6,46 | 0,11 | 11,77 | 1,18 | 100,0 |
| Пино белый № 32 | 7,26 | 0,10 | 10,86 | 0,38 | 99,8 |
| Верхнее добухтовое расстояние, см | | | | | |
| Пино белый | 7,35 | 0,13 | 16,75 | | |
| Пино черный 3 07 | 5,99 | 0,20 | 27,96 | 1,36 | 100,0 |
| Пинок белый | 5,68 | 0,16 | 20,05 | 1,67 | 100,0 |
| Пино белый № 32 | 7,59 | 0,10 | 10,12 | 0,24 | 93,1 |
| Нижнее добухтовое расстояние, см | | | | | |
| Пино белый | 5,93 | 0,10 | 16,57 | | |
| Пино черный № 07 | 5,49 | 0,15 | 21,95 | 0,44 | 99,4 |
| Пинок белый | 5,25 | 0,09 | 11,95 | 0,68 | 100,0 |
| Пино белый № 32 | 6,28 | 0,10 | 10,38 | 0,35 | 99,5 |
| Угол α , градусы | | | | | |
| Пино белый | 42,65 | 0,58 | 12,86 | | |
| Пино черный № 07 | 49,81 | 0,90 | 15,06 | 7,00 | 100,0 |
| Пинок белый | 47,85 | 0,95 | 13,77 | 5,20 | 100,0 |
| Пино белый № 32 | 44,41 | 0,85 | 14,65 | 1,76 | 95,4 |
| Угол β , градусы | | | | | |
| Пино белый | 93,69 | 0,87 | 8,71 | | |
| Пино черный № 07 | 105,26 | 0,96 | 7,59 | 11,57 | 100,0 |
| Пинок белый | 100,83 | 1,38 | 9,51 | 7,14 | 100,0 |
| Пино белый № 32 | 95,54 | 1,26 | 10,17 | 1,85 | 95,4 |

*Прим.: X – среднее арифметическое, Sx – ошибка среднего, CV – коэффициент вариации, d – разность средних, P_b – уровень вероятности разности по бутстрепу.

Из таблицы 1 видно, что по изученным показателям клоны сортогруппы Пино отличались от контрольного сорта Пино белый: степень различий варьировала от 0,05 см у клона Пино черный № 07 до 1,7 см – у Пинока белого по всем линейным признакам листьев. По угловым признакам варьирование достигало от 1,80 (Пино белый № 32) до 11,60 (Пино черный № 07).

Для оценки уровня существенности разности нами использован метод бутстрепа: он показал, что из 30 возможных 23 разности достоверны (76,7%) на 5%-ном уровне значимости ($P < 5\%$) и 20 разностей достоверны (66,7%) на 1%-ном уровне значимости ($P < 1\%$). Исключения в популяции Пино составили:

- по длине листовой пластинки у клонов Пино черный № 07 и Пино белый № 32;
- по ширине листовой пластинки у клона Пино черный № 07;
- по длине черешка у клона Пино черный № 07;
- по длине срединной жилки у клона Пино черный № 07;
- по верхнему добухтовому расстоянию у клона Пино белый № 32;
- по длине верхней боковой жилки у клона Пино белый № 32.

Из данных табл. 1 видно, что только один клон (Пинок белый) из трех по всем 10 морфометрическим признакам оказался похожим на исходный сорт, два другие клона (Пино белый № 32 и Пино черный № 07) свидетельствуют о полиморфизме сортогруппы Пино.

Как известно, сорт Рислинг как старинный европейский сорт немецкого происхождения изобилует массой полиморфных генотипов, образующих совокупности гетерогенных вариаций, из которых обычно выделяют клоны. Результаты исследований полиморфизма этой сортогруппы приведены в табл. 2.

Таблица 2. - Анализ морфологических признаков листьев исследуемых генотипов винограда в сортогруппе Рислинг

| Генотипы | *X | *S _x | *CV, % | *d | *Pb, % |
|--|-------|-----------------|--------|------|--------|
| Длина листовой пластинки, см | | | | | |
| Рислинг | 15,11 | 0,25 | 8,73 | | |
| Рислиналк | 15,36 | 0,15 | 9,23 | 0,25 | 80,7 |
| Рислинг Алькадар № 34А | 15,50 | 0,14 | 7,85 | 0,39 | 91,7 |
| Рислинг Алькадар № 34Б | 15,74 | 0,15 | 6,00 | 0,63 | 98,6 |
| Рислинг № 4-9-2 | 15,63 | 0,25 | 8,86 | 0,52 | 93,1 |
| Рислинг анапский | 15,68 | 0,39 | 10,56 | 0,57 | 89,3 |
| Ширина листовой пластинки, см | | | | | |
| Рислинг | 15,16 | 0,26 | 9,09 | | |
| Рислиналк | 14,97 | 0,16 | 9,61 | 0,19 | 73,8 |
| Рислинг Алькадар № 34А | 15,18 | 0,12 | 7,06 | 0,02 | 52,9 |
| Рислинг Алькадар № 34Б | 15,51 | 0,15 | 6,24 | 0,35 | 88,3 |
| Рислинг № 4-9-2 | 15,53 | 0,23 | 8,28 | 0,37 | 86,2 |
| Рислинг анапский | 15,43 | 0,36 | 9,96 | 0,27 | 73,5 |
| Длина черешка, см | | | | | |
| Рислинг | 8,45 | 0,16 | 10,12 | | |
| Рислиналк | 8,92 | 0,15 | 15,83 | 0,47 | 98,5 |
| Рислинг Алькадар № 34А | 8,11 | 0,13 | 13,98 | 0,34 | 95,9 |
| Рислинг Алькадар № 34Б | 8,65 | 0,15 | 11,05 | 0,20 | 83,2 |
| Рислинг № 4-9-2 | 8,74 | 0,21 | 13,13 | 0,29 | 87,2 |
| Рислинг анапский | 8,79 | 0,28 | 13,75 | 0,34 | 86,1 |
| Длина срединной жилки, см | | | | | |
| Рислинг | 9,81 | 0,19 | 10,51 | | |
| Рислиналк | 10,15 | 0,11 | 9,95 | 0,34 | 93,8 |
| Рислинг Алькадар № 34А | 10,08 | 0,10 | 8,61 | 0,27 | 89,5 |
| Рислинг Алькадар № 34Б | 10,19 | 0,11 | 6,59 | 0,38 | 95,8 |
| Рислинг № 4-9-2 | 10,10 | 0,20 | 10,73 | 0,39 | 85,1 |
| Рислинг анапский | 10,18 | 0,24 | 10,10 | 0,37 | 88,7 |
| Расстояние от верхнего бокового выступа до черешковой выемки, см | | | | | |
| Рислинг | 8,95 | 0,18 | 10,71 | | |
| Рислиналк | 8,80 | 0,10 | 10,37 | 0,15 | 76,9 |
| Рислинг Алькадар № 34А | 8,95 | 0,09 | 9,23 | 0,20 | 50,9 |
| Рислинг Алькадар № 34Б | 9,17 | 0,10 | 6,75 | 0,22 | 85,4 |
| Рислинг № 4-9-2 | 9,15 | 0,14 | 8,43 | 0,20 | 81,3 |
| Рислинг анапский | 9,21 | 0,22 | 10,07 | 0,26 | 83,1 |
| Расстояние от нижнего бокового выступа до черешковой выемки, см | | | | | |
| Рислинг | 6,90 | 0,13 | 9,82 | | |
| Рислиналк | 6,50 | 0,12 | 16,35 | 0,40 | 99,2 |
| Рислинг Алькадар № 34А | 6,82 | 0,08 | 10,60 | 0,08 | 70,3 |
| Рислинг Алькадар № 34Б | 7,04 | 0,10 | 9,29 | 0,14 | 81,5 |
| Рислинг № 4-9-2 | 6,80 | 0,12 | 9,45 | 0,10 | 71,6 |
| Рислинг анапский | 6,88 | 0,19 | 11,79 | 0,02 | 52,9 |
| Верхнее добухтовое расстояние, см | | | | | |
| Рислинг | 5,38 | 0,29 | 28,01 | | |

| | | | | | |
|----------------------------------|--------|------|-------|------|-------|
| Рислиналк | 5,04 | 0,13 | 25,21 | 0,34 | 86,1 |
| Рислинг Алькадар № 34А | 5,70 | 0,15 | 24,41 | 0,32 | 83,9 |
| Рислинг Алькадар № 34Б | 5,58 | 0,20 | 23,05 | 0,20 | 71,0 |
| Рислинг № 4-9-2 | 5,47 | 0,27 | 27,45 | 0,09 | 58,7 |
| Рислинг анапский | 5,67 | 0,33 | 24,51 | 0,29 | 74,7 |
| Нижнее добухтовое расстояние, см | | | | | |
| Рислинг | 4,37 | 0,17 | 21,17 | | |
| Рислиналк | 4,50 | 0,12 | 23,98 | 0,13 | 73,5 |
| Рислинг Алькадар № 34А | 5,12 | 0,10 | 18,12 | 0,75 | 100,0 |
| Рислинг Алькадар № 34Б | 5,04 | 0,14 | 18,12 | 0,67 | 99,8 |
| Рислинг № 4-9-2 | 5,11 | 0,16 | 17,46 | 0,74 | 99,9 |
| Рислинг анапский | 5,00 | 0,21 | 17,70 | 0,63 | 99,0 |
| Угол α , градусы | | | | | |
| Рислинг | 48,25 | 1,20 | 13,10 | | |
| Рислиналк | 48,07 | 0,99 | 18,97 | 0,18 | 54,9 |
| Рислинг Алькадар № 34А | 49,67 | 0,89 | 16,05 | 1,42 | 83,2 |
| Рислинг Алькадар № 34Б | 50,43 | 1,09 | 13,70 | 2,18 | 91,5 |
| Рислинг № 4-9-2 | 50,74 | 1,58 | 17,36 | 2,49 | 90,0 |
| Рислинг анапский | 45,78 | 1,58 | 14,65 | 2,47 | 90,0 |
| Угол β , градусы | | | | | |
| Рислинг | 98,32 | 1,84 | 9,90 | | |
| Рислиналк | 103,86 | 1,23 | 10,95 | 5,54 | 99,5 |
| Рислинг Алькадар № 34А | 104,41 | 0,98 | 8,46 | 6,09 | 99,9 |
| Рислинг Алькадар № 34Б | 104,20 | 1,04 | 6,34 | 5,88 | 99,8 |
| Рислинг № 4-9-2 | 106,06 | 2,21 | 11,60 | 7,74 | 99,8 |
| Рислинг анапский | 97,72 | 2,22 | 9,65 | 0,60 | 57,7 |

*Примечание: \bar{X} – среднее арифметическое, Sx – ошибка среднего арифметического, CV – коэффициент вариации, d – разность средних, Pb – уровень вероятности разности по бутстрепу.

В сортогруппе Рислинг (табл. 2) изучаемые клоны по морфологии листьев отличались от контрольного сорта по 13 из 50 (26%) наблюдений при $P < 5\%$ и по 9 (18%) - при $P < 1\%$. Линейные различия варьировали от 0,02 см (Рислинг анапский) до 0,75 см (Рислинг Алькадар № 34А), угловых – от 0,180 (Рислиналк) до 7,740 (Рислинг № 4-9-2).

С помощью метода бутстрепа в сортогруппе Рислинг выявлены достоверные различия между контролем и клонами по следующим признакам:

- по длине листовой пластинки н клона Рислинг Алькадар № 34Б;

- по длине черешка у клонов Рислиналк и Рислинг Алькадар № 34А;

- по длине срединной жилки у клона Рислинг Алькадар № 34Б;

- по длине нижней боковой жилки у клона Рислиналк;

- по нижнему добухтовому расстоянию у клонов Рислинг Алькадар № 34А, Рислинг Алькадар № 34Б, Рислинг № 4-9-2 и Рислинг анапский;

- по углу β у клонов Рислиналк, Рислинг Алькадар № 34А, Рислинг Алькадар № 34Б и Рислинг № 4-9-2.

Коэффициенты вариации признаков сортогруппы Пино оказались относительно невысокими и составляли от 7% до 28%, что свидетельствует о хорошем уровне выравненности морфометрических данных у каждого исследуемого клона. В сортогруппе Рислинг он находился в тех же пределах, 6–28%.

Биометрико-ампелографическое изучение морфометрических признаков обеих популяций винограда показало, что выявлено значительное отличие почти всех клонов винограда от исходных сортов по разнообразным показателям листа, за исключением клона Пинок белый, который по всем параметрам листа оказался идентичным контрольному сорту Пино белый.

Таким образом, поскольку по большинству морфометрических признаков листьев изученные генотипы в обеих сортогруппах Пино и Рислинг биометрически различаются, это свидетельствует о наличии их фенотипического полиморфизма. Таким образом, все генотипы фенотипически отличаются не только от исходных форм, кроме одного Пинока белого, но и различаются между собой.

С использованием метода системно-когнитивного анализа [2] морфометрических данных листьев обеих исследованных сортогрупп Пино и Рислинг (по их 10 признакам) было обнаружено следующее сходство генотипов (табл. 3).

Таблица 3. - Результаты системно-когнитивного анализа морфометрических данных сортогруппы Пино

| № | Генотип | Пино белый | Пино белый № 32 | Пино черный № 07 | Пинок белый |
|---|------------------|------------|-----------------|------------------|-------------|
| 1 | Пино белый | 100,00000 | 34,60015 | -46,40238 | -45,05768 |
| 2 | Пино белый № 32 | 34,60015 | 100,00000 | -36,76949 | -15,44583 |
| 3 | Пино черный № 07 | -46,40238 | -36,76949 | 100,00000 | 14,75945 |
| 4 | Пинок белый | -45,05768 | -15,44583 | 14,75945 | 100,00000 |

Как видно из таблицы 3, клон Пино белый № 32 по всей системе (комплексу) признаков листа больше всего похож на контрольный сорт Пино белый, а на остальные клоны не похож; клон Пино черный № 07 по всей системе признаков листа больше всего похож на Пинок белый, как и Пинок белый - на клон Пино черный № 07.

Таблица 4. - Результаты системно-когнитивного анализа морфометрических данных сортогруппы Рислинг

| № | Генотип | Рислиналк | Рислинг | Рислинг № 4-9-2 | Рислинг Алькадар № 34А | Рислинг Алькадар № 34Б | Рислинг анапский |
|---|------------------------|-----------|----------|-----------------|------------------------|------------------------|------------------|
| 1 | Рислиналк | 100,00 | 4,26 | -21,65 | -22,71 | -20,59 | -15,86 |
| 2 | Рислинг | 4,26 | 100,00 | -0,12 | -25,25 | -10,29 | 10,46 |
| 3 | Рислинг № 4-9-2 | -21,64 | -0,13579 | 100,00 | -2,07 | 34,03 | 38,79 |
| 4 | Рислинг Алькадар № 34А | -22,71 | -25,25 | -2,07 | 100,00 | 0,57 | -20,26 |
| 5 | Рислинг Алькадар № 34Б | -20,59 | -10,29 | 34,03 | 0,57 | 100,00 | 9,27 |
| 6 | Рислинг анапский | -15,86 | 10,46 | 38,79 | -20,26 | 9,27 | 100,00 |

Как видно из таблицы 4, клон Рислиналк по всей системе признаков листа больше всего похож на Рислинг, а на остальные не похож; Рислинг анапский по всей системе признаков листа больше всего похож на Рислинг № 4-9-2, меньше – на Рислинг и Рислинг Алькадар № 34Б, а на остальные не похож.

В заключение отметим, что впервые для многопараметрической типизации, многопараметрической (системной) идентификации и кластерного анализа сортов винограда по морфометрическим признакам листьев был применен метод искусственного интеллекта «Системно-когнитивный анализ» (СК-анализ) и его инструментарий – система «Эйдос» [2], которые предоставляют многочисленные и разнообразные возможности. Последние освещать в данной статье, по нашему мнению, нецелесообразно. Отметим лишь, что он, в частности, позволяет:

1. Провести *многопараметрическую типизацию*, т.е. непосредственно на основе описаний конкретных объектов, относящихся к тем или иным обобщенным категориям (классам) сформировать обобщенные образы классов, в которых признаки ранжированы в порядке убывания степени характерности для данного класса (т.е. в порядке убывания количества информации в признаке о принадлежности к данному классу). При этом для каждого из классов определяется наиболее характерный для него признак.

2. Непосредственно на основе эмпирических данных рассчитать, какое количество информации содержится в каждом из признаков (градации описательной шкалы, которая может быть номинальной, порядковой или отношений) о принадлежности обладающего этим признаком объекта к каждой из обобщенных категорий (классов).

3. Провести *многопараметрическую (системную) идентификацию*, т.е. количественно сравнивать конкретные объекты с обобщенными образами классов по всей совокупности (системе) признаков. При этом объект (согласно лемме Неймана-Пирсона) считается относящимся к тому классу, о принадлежности к которому в его *системе признаков* содержится наибольшее количество информации.

4. Сравнить обобщенные образы классов друг с другом по всей совокупности (системе) признаков и сформировать группы наиболее сходных классов (*кластерный анализ*).

5. Сформировать конструкты классов, т.е. системы наиболее непохожих друг на друга кластеров (*конструктивный анализ*).

6. Разделить классы с высокой вариабельностью признаков входящих в них объектов на подклассы с минимальной внутренней вариабельностью и сформировать NCD-дерево, отражающее порядок их разделения.

Выводы. Использование многомерных методов биометрии, в том числе системно-когнитивного анализа, позволило обосновать полиморфизм популяций винограда Пино и Рислинг по морфометрическим признакам листа и убедиться в наличии разных по продуктивности клонов. Выделенные высокопродуктивные клоны этих популяций тестированы молекулярно-генетическим методом, что подтвердило обоснованность наличия полиморфизма по количественным признакам морфометрии листа. Все использованные методы дали результаты, хорошо согласующиеся друг с другом, что повышает достоверность сделанных на их основе выводов.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Ампеелография СССР. - М.: Пищепромиздат, 1946-1984. – Т. 1-11.
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. Островерхов В.О., Трошин Л.П. Методические рекомендации по оценке стабильности количественных признаков у сортов винограда / ВНИИВиПП "Магарач". - Ялта, 1986. - 86 с.
4. Подваленко П.П., Звягин А.С., Трошин Л.П. Клоновая селекция – современная основа подъема продуктивности виноградников // Научный журнал КубГАУ. – 2009. - № 51 (07). – 25 с. <http://ej.kubagro.ru/2009/07/>.
5. Трошин Л.П. Ампеелография и селекция винограда. – Краснодар: РИЦ «Вольные мастера», 1999. – 138 с.: цв. вкладка.
6. Трошин Л.П., Животовский Л.А. Методические рекомендации по клоновой селекции винограда на продуктивность / ВНИИВиПП "Магарач". Ин-т общей генетики им. Н.И.Вавилова. - Ялта, 1987. - 36 с.

7. Трошин Л.П., Звягин А.С., Подваленко П.П. Анализ генетического разнообразия клонов сортогрупп Пино и Рислинг с использованием микросателлитных маркеров // Материалы XVIII Международного научного симпозиума «Нетрадиционное растениеводство. Селекция и генетика. Эниология. Экология и здоровье». 17-26 сентября 2009 г. – Симферополь, 2009. – С. 308-313.

8. Трошин Л.П., Симонова Н.Л. Сортимент виноградных насаждений России продолжает совершенствоваться // Труды КубГАУ. – Краснодар, 2009. - №4 (19). – С. 72-76.

9. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. – М.: ИНФРА-М, 2003. – 544 с.

10. Энциклопедия виноградарства. - Кишинев: МСЭ, 1986-1987. – Т. 1-3.



Пино белый



Рислинг рейнский

ОБЩИЕ ВЫВОДЫ

Среди современных методов НИР самым эффективным, пока малораспространенным, является «Эйдос», удививших мир биометриков, научных сотрудников в разных областях всего живого на Земле.

За ним, по нашему глубокому убеждению, большое будущее!

1. Значимость полиморфизма сортов

Полиморфизм сортов винограда, отражающий одновременное существование нескольких фенотипически и генотипически отличающихся групп растений, имеет важное значение для виноградарства. Клоны одного сорта обеспечивают возможность формирования гармоничного ампелоценоза, повышающего адаптивность и экономическую эффективность производства. Применение системного подхода к изучению полиморфизма позволяет лучше понять механизмы формирования сортов и их адаптационных способностей.

2. Методология анализа

Автоматизированный системно-когнитивный анализ (АСК-анализ), основанный на системной теории информации, представляет собой эффективный инструмент для изучения морфологических особенностей листьев винограда. Использование интеллектуальной системы «Эйдос» позволяет оцифровать изображения листьев, создать их математические модели и сформировать обобщенные образы, что существенно упрощает процесс идентификации сортов и клонов.

3. Инновации в оцифровке листьев

Предложенный подход к оцифровке изображений листьев на основе полярной системы координат, стандартизации положения, размера и угла поворота изображений позволяет получить инвариантные результаты. Это дает возможность

минимизировать влияние внешних факторов, таких как шум и случайные искажения формы листа, что повышает точность анализа.

4. Форма листа как источник информации

Контур листа рассматривается как информационное сообщение, содержащее сведения об истинной форме сорта или клона (сигнал) и случайные отклонения (шум), обусловленные воздействием внешней среды. Интеллектуальная система «Эйдос» эффективно выделяет сигнал, подавляя шум, что позволяет формировать единый образ формы листа для каждого сорта или клона.

5. Применение в ампелографии

Программный инструментарий «Эйдос» обеспечивает:

- количественное определение сходства и различий между сортами и клонами винограда;
- верификацию моделей путем идентификации конкретных листьев с обобщенными образами сортов или клонов;
- многопараметрическую типизацию, позволяющую глубже изучить морфологическое разнообразие.

Эти возможности открывают новые перспективы для исследований в ампелографии и внедрения инновационных технологий в практическое виноградарство.

6. Заключение

Таким образом, использование методов АСК-анализа и системы «Эйдос» для изучения полиморфизма винограда является перспективным направлением исследований. Это способствует углубленному пониманию механизмов формирования сортов, повышению точности идентификации и разработке рекомендаций по оптимизации сортового состава для достижения высокой урожайности и качества продукции.

Научное издание

Луценко Евгений Вениаминович
Трошин Леонид Петрович

**КОГНИТИВНОЕ МОДЕЛИРОВАНИЕ
В АМПЕЛОГРАФИИ**

Монография

В авторской редакции
Компьютерная верстка – Е. В. Луценко
Обложка – Е. В. Луценко

Подписано в печать 12.12.24. Формат 70 × 100 ¹/₁₆
Усл. печ. л. – 9,688. Уч.-изд. л. – 6,458.
Тираж 100 экз. Заказ №

Типография Принтстори
www.printstory.ru