

Интернет-журнал «Наукоедение» ISSN 2223-5167 <http://naukovedenie.ru/>

Том 8, №2 (2016) <http://naukovedenie.ru/index.php?p=vol8-2>

URL статьи: <http://naukovedenie.ru/PDF/18TVN216.pdf>

DOI: 10.15862/18TVN216 (<http://dx.doi.org/10.15862/18TVN216>)

Статья опубликована 25.03.2016.

**Ссылка для цитирования этой статьи:**

Джевага К.А., Мезенцев А.О. Функциональная модель процесса когнитивно-образной визуализации разнородных данных // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 8, №2 (2016) <http://naukovedenie.ru/PDF/18TVN216.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ. DOI: 10.15862/18TVN216

**УДК 004.92, 004.891**

**Джевага Константин Александрович**

ГКОУ ВПО «Академия Федеральной службы охраны Российской Федерации», Россия, Орёл<sup>1</sup>  
Сотрудник  
Кандидат технических наук  
E-mail: [dka06@mail.ru](mailto:dka06@mail.ru)

**Мезенцев Александр Олегович**

ГКОУ ВПО «Академия Федеральной службы охраны Российской Федерации», Россия, Орёл  
Сотрудник  
E-mail: [aomezentsev@gmail.com](mailto:aomezentsev@gmail.com)

## **Функциональная модель процесса когнитивно-образной визуализации разнородных данных**

**Аннотация.** В статье рассматривается вариант решения задачи аналитической обработки больших объёмов разнородных исходных данных, используя возможности человека по визуальному анализу данных, являющихся предметом исследования когнитивной компьютерной графики, предложена функциональная модель данного процесса визуализации информации. Актуальность задачи заключается в развитии концепции «открытых данных», быстром росте объёма анализируемых данных, необходимости поиска закономерностей, получения нового знания, отсутствии научно-обоснованных способов синтеза адекватного исходным данным графического образа, учитывающие индивидуальные особенности восприятия лица, принимающего решения. Использование возможностей человека по визуальному анализу данных, являющихся предметом исследования когнитивной компьютерной графики, позволяет повысить эффективность поиска фактов наличия закономерностей, ограничения исходного объёма данных, принятия решения о целесообразности дальнейшего анализа при решении задачи и пр. Предлагается подход к формированию графического образа, пригодного для использования в системах анализа больших объёмов данных. Описаны факторы, влияющие на процесс формирования графического образа. Предложена общая модель графического образа, перечислены предъявляемые к нему требования. Разработанная на его основе функциональная модель процесса когнитивно-образной визуализации таких данных. Наличие в функциональной модели обратной связи позволяет учесть факторы, связанные с психофизиологическими

---

<sup>1</sup> 302034, г. Орёл, ул. Приборостроительная, д. 35

особенностями восприятия информации и физическими особенностями окружения системы отображения информации.

**Ключевые слова:** открытые данные; большие данные; обработка больших объемов данных; интеллектуальный анализ данных; компьютерная графика; когнитивная компьютерная графика; аналитическая обработка данных; графический образ; визуализация информации

## Введение

Одной из важных и наиболее ресурсоемких задач процесса контроля состояния и динамики изменения объектов государственного управления является периодический поиск и анализ большого объема информации, содержащейся в открытых источниках в различных формах представления, необходимой для принятия решений. В большинстве случаев данная задача решается в рамках развития концепции «открытых данных», предполагающей доступ посредством сети «Интернет» к открытым данным, содержащимся в информационных системах органов государственной власти Российской Федерации. В Российской Федерации область открытых данных развивается после публикации Указа Президента Российской Федерации от 7 мая 2012 г. № 601<sup>2</sup>, предполагающего обеспечить доступ в сети «Интернет» к открытым данным, содержащимся в информационных системах органов государственной власти Российской Федерации.

Значительный объем исходных данных и многообразие их типов и форм представления с одной стороны и необходимость разностороннего описания объектов управления с учетом их взаимосвязей и особенностей функционирования с другой, породили проблему автоматизированного анализа открытых данных с целью выявления закономерностей, получения новых знаний об объекте управления, восполнения недостающих данных, прогнозирования поведения объекта управления и т.д.

Решение данной проблемы находится в компетенции отрасли информационной науки (information science), обозначаемой термином интеллектуальный анализ данных (ИАД, data mining) [1]. Частичное решение этой проблемы достигнуто за счет разработки и реализации технологий кластерного анализа [2], построения ассоциативных деревьев [3], регрессионного анализа [4], анализа временных рядов [5] и др. Однако решающая роль в интерпретации полученных результатов и принятии решений отводится человеку, рассматриваемому в качестве активного компонента системы аналитической обработки информации.

В связи с этим, немаловажную роль играют способы и формы представления, как исходных данных (ИД), так и результатов аналитической обработки, учитывающие особенности восприятия и мышления.

Важным этапом ИАД является подготовка данных, заключающаяся в очистке ИД от «информационного шума», неполных или недостоверных значений. При этом возникает опасность искажения информационного описания объекта управления.

Использование возможностей человека по визуальному анализу информации, являющихся предметом исследования когнитивной компьютерной графики (ККГ), при подготовке данных позволит повысить эффективность этого этапа в части поиска фактов

---

<sup>2</sup> Об основных направлениях совершенствования системы государственного управления [Текст]: Указ Президента Российской Федерации от 07.05.2012 № 601 // Российская газета. 09.05.2012. № 5775. – Режим доступа: <http://www.rg.ru/2012/05/09/gosupravlenie-dok.html>, свободный.

наличия закономерностей, ограничения исходного объёма ИД, принятия решения о целесообразности дальнейшего анализа при решении задачи и т.д. Тема применения когнитивных графических образов при анализе данных получила в последнее время широкую известность [6, 7]. Научная школа Горохова В.Л. [8, 9], исследует проблемы использования когнитивных технологий визуализации данных в системах поддержки принятия решений.

Вопросам анализа и визуализации больших объемов данных (BigData) посвящено множество работ отечественных [10, 11, 12] и зарубежных [13, 14] авторов. Вышеперечисленные работы рассматривают как теоретические, так и практические аспекты применения ККГ для визуального анализа информации. Важной их особенностью является определение типа графического образа (ГО) с помощью мнений экспертов, эмпирическим или директивным путём, а так же использование ГО на этапе принятия решения, после обработки, ручного или автоматизированного поиска закономерностей информации.

При этом, в данных работах не рассматриваются научно-обоснованные способы синтеза адекватного исходным данным графического образа, учитывающие индивидуальные особенности восприятия, а также факторы среды (окружения) в которой происходит наблюдение визуального образа данных (освещенность рабочего места, тип поверхности отображения – самосветящаяся, несамосветящаяся, размер поверхности отображения – индивидуальный просмотр, общего пользования и др.).

В данной работе предлагается подход к формированию графического образа, пригодного для использования в системах анализа больших объемов данных и разработанная на его основе функциональная модель процесса когнитивно-образной визуализации таких данных.

### **Анализ исходных данных**

Во исполнение Указа Президента Российской Федерации от 7 мая 2012 г. № 601, была разработана и одобрена Концепция размещения государственными органами и органами местного самоуправления информации о своей деятельности в форме открытых данных. Согласно положениям концепции<sup>3</sup>, открытые данные рассматриваются как один из наиболее удобных форматов для поиска, обработки и дальнейшего использования открытой, общедоступной и достоверной информации, способствующих реализации принципа информационной открытости при реализации государственных полномочий и функций.

На общероссийском Портале открытых данных<sup>4</sup>, а также на порталах, созданных некоторыми отдельными регионами и городами, обеспечен доступ к 5293 базам данных (по состоянию на февраль 2016 года), организованных по 16 основным темам, среди которых – метеоданные, образование, транспорт и здоровье.

Основными принципами открытых данных являются<sup>5</sup>:

- первичность;
- полнота;

---

<sup>3</sup> Концепция открытости федеральных органов исполнительной власти [Текст]: Распоряжение Правительства РФ от 30 января 2014 г. № 93-р // СЗ РФ. 2014. №5. Ст. 547.

<sup>4</sup> Портал открытых данных Российской Федерации [Электронный ресурс] – Режим доступа: <http://data.gov.ru/>, свободный.

<sup>5</sup> Что такое Открытые данные [Электронный ресурс] – Режим доступа: <http://opengovdata.ru/definition/>, свободный.

- актуальность;
- пригодность к машинной обработке;
- отсутствие дискриминации по доступу;
- отсутствие неизвестных форматов;
- лицензионная чистота.

Предложим следующую классификацию открытых данных, содержащихся в перечисленных источниках:

- по форме представления:
  - текстовые данные;
  - табличные данные;
  - числовые данные текстом и т.д.;
- по способу описания характеристик:
  - количественные данные;
  - качественные данные;
- по виду данных:
  - статистические данные по различной тематике;
  - данные об объектах реального мира;
  - фактографические данные.

Из анализа результатов исследований корпорации EMC, имеющей продолжительный опыт работы с большими данными, хранимыми в облачных системах<sup>6</sup>, следует, что объём данных информационной вселенной (Интернет) растёт по закону, близкому к экспоненциальному (2012 г. – 2,8 трлн. гигабайт, 2014 г. – 4,4 трлн. гигабайт), при этом количество неструктурированных, слабоформализуемых многомерных данных составляет более 90%. Полагая, что закономерности в поведении системы в целом справедливы и для ее частей, будем считать, что рост объёма анализируемых данных, содержащихся в приведенных источниках в различных формах представления, также происходит по закону, близкому к экспоненциальному.

Снизить ресурсоемкость обработки данных позволяет полная или частичная автоматизация процессов сбора, первичной обработки и анализа данных. При этом многообразие форм представления информации, неструктурированное или слабоструктурированное описание объектов, событий, явлений, избыточность или, напротив, неполнота описания затрудняют использование стандартных методов. Активное использование когнитивных и аналитических возможностей человека в процессе обработки данных требует их представления в формах, удобных для восприятия, исключающих неоднозначность интерпретации, стимулирующих мыслительную деятельность.

---

<sup>6</sup> The Digital Universe of Opportunities, [Электронный ресурс] Корпорация EMC, 2014 г. – Режим доступа: <http://russia.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>, свободный.

## Особенности визуального представления разнородных данных

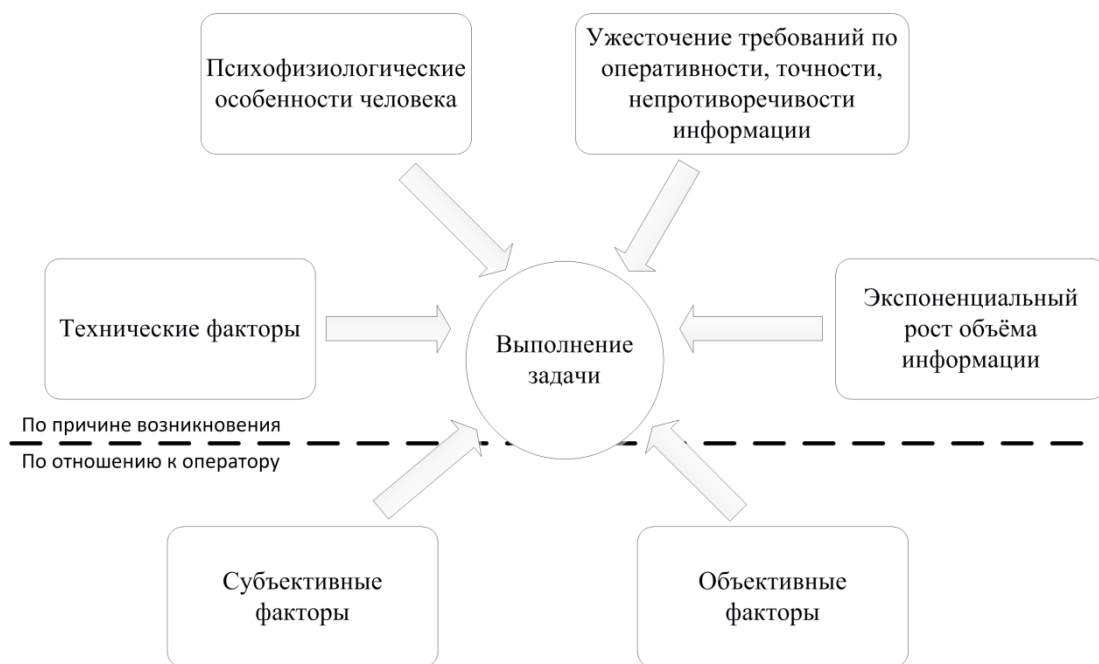
Задача визуализации числовых и хорошо структурированных данных представляется достаточно тривиальной и хорошо отработанной. Визуализация нечисловых (качественных) и смешанных данных не имеет единого подхода. Используемые при построении графического образа традиционные метафоры, вызывая у эксперта ассоциации с исследуемыми характеристиками, не позволяют в большинстве случаев выявить скрытые закономерности и прочую производную информацию об исследуемом объекте. Частный случай применения графических образов для отображения информации описан в работе А.В. Толока [10]. Одним из путей решения задачи визуализации и анализа большого объема слабоформализованных смешанных данных является применение методов ККГ [15].

Использование ККГ заключается в достижении "когнитивного" эффекта – получения нового знания, для которого необходимо использование множества абстрактных форм, не вызывающих прямых ассоциаций у лиц, принимающих решение (ЛПР), стимулируя их мыслить более общими формами. Проблемой является выбор такой формы графического образа, которая позволит возникнуть необходимому "когнитивному эффекту" получения знания.

Реальные технические системы, использующие технологию когнитивной графики, существуют и разрабатываются в ИПС РАН, ИСА РАН, МЭИ. В медицине – для мониторинга состояния больных, в технике – для обслуживания запуска ракет, стыковки космических аппаратов, контроля состояния реакторов АЭС и др. приложениях. При этом «объект управления» и его отображение в когнитивный графический образ определен и понятен для оператора. Например, для визуальной оценки состояния энергоблока АЭС на экран выводятся динамический образ сканируемых элементов, цвет которых отражает близость параметров к критическим порогам [16].

При разработке методологии формирования ГО следует учесть ряд факторов, влияющих на его качество, выражающееся легкостью и однозначностью интерпретации содержащейся в нем информации. Предлагается классификация этих факторов по следующим признакам (рисунок 1):

- по отношению к оператору и подверженности автоматизации:
  - объективные, не зависящие от оператора, влияние которых можно снизить с помощью автоматизации соответствующих процессов;
  - субъективные, зависящие от оператора, частично подверженные автоматизации;
- по причине возникновения:
  - обусловленные психофизиологическими особенностями оператора при восприятии информации;
  - обусловленные ужесточением требований по оперативности, точности, непротиворечивости поступающей информации;
  - экспоненциальный рост объема информации;
  - технические факторы, к которым относятся ограничения средств отображения по цветовому диапазону, разрешению, размерам изображения.



**Рисунок 1.** Классификация факторов, влияющих на процесс представления информации

### Общая модель графического образа

Понятие графического образа в ККГ может быть определено как визуальное представление предметов, явлений, процессов и их окружения, отражающее их значимые свойства, синтезированное ЭВМ на основе математических моделей различных видов [15].

Решение задачи синтеза ГО анализируемой информации затрудняется отсутствием унифицированного подхода к его формированию, учитывающего разнородность исходных данных, особенности восприятия пользователем визуальной информации, и другие факторы и ограничения, описанные выше.

Формально ГО представим кортежем (1):

$$O = \langle G, E, T \rangle, \quad (1)$$

где:

$O$  – графический образ;

$G$  – множество пространственно-геометрических характеристик;

$E$  – множество энергетических характеристик;

$T$  – множество динамических характеристик.

Важное значение имеет информационная мерность ГО – количество изменяющихся параметров данного ГО, передающих своими изменениями информацию о контролируемых параметрах объектов реального мира.

Говоря о динамических характеристиках, имеем в виду, прежде всего, изменение пространственно-геометрических и энергетических характеристик ГО. Тогда ГО можно представить динамической моделью (2):

$$O(t) = \langle \{g_1(t), g_2(t), \dots, g_i(t)\}, \{e_1(t), e_2(t), \dots, e_j(t)\} \rangle, \quad (2)$$

где:

$t$  – момент времени;

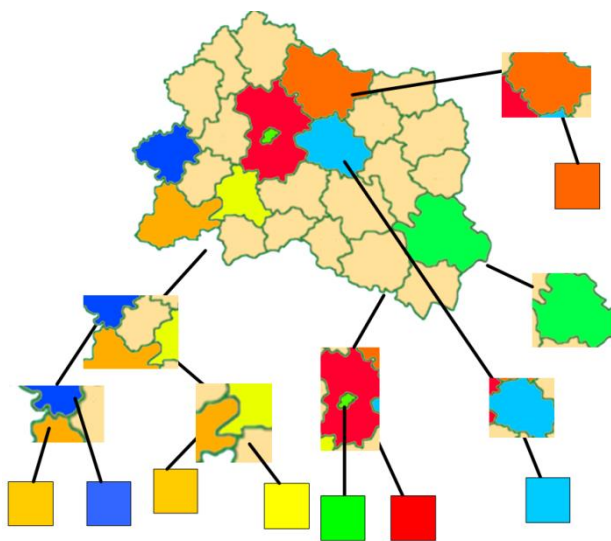
$g_i(t)$  – значение  $i$ -той пространственно-геометрической характеристики в момент времени  $t$ ;

$e_i(t)$  – значение  $i$ -той энергетической характеристики в момент времени  $t$ .

Особенностями предлагаемого подхода являются:

1. Определение, перед созданием ГО, множества "контролируемых параметров", как количественных, так и качественных, значения которых могут быть получены в автоматизированном порядке.
2. Накопление библиотеки когнитивных примитивов в соответствии с предметной областью и психофизиологическими особенностями ЛПР.
3. Автоматизированная предобработка исходных данных (например, многослойной нейронной сетью) с целью кластеризации значений контролируемых параметров.
4. Представления кластеризованных значений контролируемых параметров в виде динамически изменяемых характеристики ГО (2), с учётом корректировок ЛПР.

На рисунке 2 ГО представлен в виде иерархической модели, уровни которой связаны как «Целое – Часть целого». Проведенные исследования показали справедливость моделей (1) и (2) для каждого уровня иерархии.



**Рисунок 2.** Абстрактные уровни формирования графического образа  
Особенности каждого уровня иерархической модели ГО отражены в таблице 1.

**Таблица 1**

**Особенности уровней формирования ГО**

Уровень	Факторы, влияющие на ГО	Изменяющиеся характеристики ГО	Физический смысл уровня	Пример
Пиксель	Технические факторы, ограничивающие изображение со стороны физических параметров средства отображения информации	Диапазон цвета точки	Качественная характеристика значения отдельного параметра	
Кластер	Технические и психофизиологические факторы	Размер кластера, положение в рамках средства отображения информации	Группа качественных характеристик значений отдельных параметров, выделенная по какому-либо признаку	
Группы кластеров	Технические и психофизиологические факторы	Взаимное расположение кластеров и его динамическое изменение в рамках средства отображения информации	Признак выделения отдельных параметров в кластер	
Графический образ	Все группы факторов	Взаимное расположение групп кластеров и его динамическое изменение в рамках средства отображения информации	Необходимый объём данных, значений параметров	

Методы и алгоритмы «кодирования» (преобразования) информации в графический образ реализуют отображение множества анализируемых данных во множество характеристик ГО.

Таким образом, используя находящиеся в информационном хранилище АС правила преобразования информации, а так же учитывая особенности объекта анализа и предпочтения ЛПР производится выбор параметров объекта и отображения их в виде различных значений характеристик ГО. Т.е. каждый необходимый для отображения (что так же определяется важностью) параметр задаёт внешний вид элемента ГО.

**Функциональная модель процесса когнитивно-образной визуализации разнородных данных**

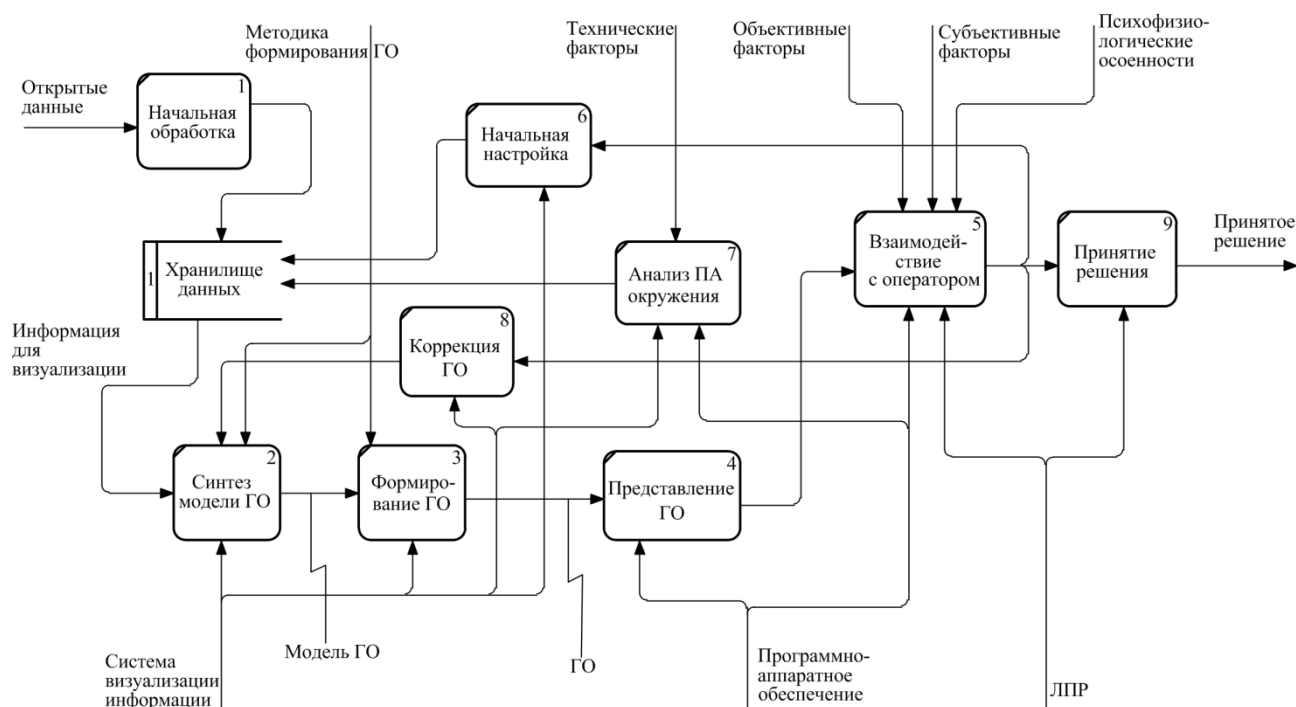
Таким образом, на основании вышеизложенного, сформулированы основные требования к ГО, выполнение которых позволит ему быть пригодным для аналитической обработки данных. ГО должен:

- сохранять значимые свойства исходных данных;
- иметь в основе математическую модель исходных данных;
- обеспечивать отсутствие влияния полноты исходных данных и статистических выбросов на восприятие оператором свойств и параметров анализируемого объекта;
- быть представленным в виде многомерного изображения;



- учитывать психофизиологические особенности пользователя;
- учитывать технологические ограничения средств вывода информации (экранов мобильных устройств, мониторов, проекторов и пр.);
- превосходить по информативности, непротиворечивости другие формы представления информации;
- учитывать экспоненциальную скорость роста исходных данных при выборе математической модели, лежащей в его основе.

С целью реализации вышеперечисленных требований разработана функциональная модель процесса адаптивной когнитивно-образной визуализации информации (рисунок 3).



**Рисунок 3.** Функциональная модель процесса адаптивной когнитивно-образной визуализации информации

Общий алгоритм функционирования представляет собой набор следующих утверждений, начальных состояний и блоков.

На этапе начальной обработки данных (1) загруженные открытые данные подвергаются систематизации и начальному анализу с целью помещения в хранилище данных. В хранилище данных так же находится набор «эталонных» данных, необходимых для первоначальной подстройки под психофизические особенности конкретного ЛПП.

В основном режим функционирования участвуют следующие блоки функциональной модели:

1. Анализ программно-аппаратного окружения средств взаимодействия с оператором и настройка соответствующих параметров ГО (7) и занесение в хранилище данных результата.
2. Синтез модели ГО на основе загруженных данных и параметров коррекции ГО (2).
3. Формирование ГО (3).

4. Вывод изображения (4) на средства отображения информации.
5. Взаимодействие с оператором с целью уточнения данных, изменения точки обзора ГО и т.д. (5).
6. Коррекция ГО в соответствии с данными, полученными в результате взаимодействия с оператором (8).

Режим первоначальной настройки отличается использованием в качестве исходных данных «эталонных» значений показателей. Наличие в разработанной структуре обратной связи (5, 8) позволяет учесть факторы, связанные с психофизиологическими особенностями восприятия информации конкретным пользователем и с физическими особенностями окружения системы отображения информации, что повышает уровень участия ЛПР в процессе принятия решения и повышает адаптивность визуализации информации.

Множества исходных данных и результатов анализа помимо детерминированных или стохастических числовых значений содержат качественные слабоформализованные данные. Элементы структуры графического образа также описываются нечеткими категориями. Эти факторы обуславливают использование при разработке подсистемы формирования ГО методов теории нечетких множеств, нечетких алгоритмов и искусственных нейронных сетей.

### Заключение

Методы когнитивной компьютерной графики, такие как цветовая маркировка таблиц, диаграммы, графики, привязка данных к картам и др., применяются в современных информационно-аналитических системах как правило на этапе представления информационно-аналитических материалов лицу, принимающему решение.

Особенностью предлагаемого подхода является использование этих методов уже на этапе подготовки данных, что позволит значительно снизить количество рутинных операций обработки больших объемов исходных данных и использовать лицо принимающее решение как активный элемент информационной системы.

Наибольшую сложность при этом представляет задача формирования ГО, отражающего как числовые, так и качественные данные. Для её решения предложена функциональная модель процесса адаптивной визуализации информации. Наличие в предлагаемом варианте обратной связи позволяет учесть влияние различных групп факторов.

Предложенный подход определяет программу формирования методологии когнитивно-образного анализа больших объёмов разнородных данных, включающей исследование свойств, разработку, модификацию и объективную оценку качества ГО.

### ЛИТЕРАТУРА

1. Барсегян, А.А. Технология анализа данных: Datamining, Visialmiming, TextMining, OLAP [Текст] / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – 2-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2007. – 384 с.
2. Ким, Дж.-О. Факторный, дискриминантный и кластерный анализ: Пер. с англ. [Текст] / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.; Под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.
3. Кнут, Д.Э. Искусство программирования, том 3. Сортировка и поиск [Текст] / Дональд Эрвин Кнут – М: Вильямс, 2007 г. – 824 с.

4. Шашков, В.Б. Прикладной регрессионный анализ. Многофакторная регрессия [Текст] / В.Б. Шашков – Учебное пособие. – Оренбург: ГОУ ВПО ОГУ, 2003. – 363 с.
5. Афанасьев, В.Н. Анализ временных рядов и прогнозирование: учебник [Текст] / В.Н. Афанасьев, М.М. Юзбашев. – Изд. 2-е, перераб. и доп. – М.: Финансы и статистика; ИНФРА-М, 2010. – 320 с.: ил.
6. Цаплин, В.В. Когнитивные технологии визуализации многомерных данных для интеллектуальной поддержки принятия решений [Текст]: В.В. Цаплин, В.В. Витковский, В.Л. Горохов // Программные продукты и системы. – 2014. – № 3 (107). – с. 22-25.
7. Манько, Н.Н. Когнитивная визуализация педагогических объектов в современных технологиях обучения [Текст]: Н.Н. Манько // Образование и наука. – 2009. – №8 (65). – с. 10-30.
8. Горохов, В.Л. Современные методы когнитивной визуализации многомерных данных [Текст] / В.Л. Горохов, А.А. Лукьянец, А.Г. Чернов. - Томск: Некоммерческий фонд развития региональной энергетики, 2007. - 216 с.: ил.
9. Горохов, В.Л. Когнитивные технологии управления рисками в техносфере. Проекты и исследования. [Текст]: В.Л. Горохов, В.В. Цаплин, В.В. Витковский // Вестник гражданских инженеров: СПб ГАСУ – 2011 – № 3 – с. 122–127.
10. Толоч, А.В. Графические образы-модели в информационных технологиях [Текст]: А.В. Толоч // Прикладная информатика. – 2009. – №4. – с. 31-40.
11. Бабурин, В.А. Технологии Big Data в сервисе: новые рынки, возможности и проблемы [Текст]: В.А. Бабурин, М.Е. Яненко // Техничко-технологические проблемы сервиса. – 2014. – №1 (27) – с. 100-105.
12. Клеменков, П.А., Большие данные: современные подходы к хранению и обработке [Текст]: П.А. Клеменков, С.Д. Кузнецов. // Труды Института системного программирования РАН. – 2012. – № 23 – с. 143-156.
13. Big Data Visualization: Turning Big Data into Big Insights. The Rise of Visualization-based Data Discovery Tools. White Paper. Intel IT Center. March 2013.
14. Shneiderman. The big picture for big data: Visualization. Science, 343:730, February 2014.
15. Компьютерная графика: методы, модели и средства преобразования графической информации. [Текст]: Том 2. Теоретические основы обработки изображений и прикладные аспекты компьютерной графики: монография в 2-х томах. / Под ред. д.т.н. А.П. Фисуна, д.т.н. И.С. Константинова, д.т.н. В.А. Минаева // Орел: «ОГУ», «ОрелГТУ», 2010. – 238 с.
16. Башлыков, А.А. Методы когнитивной графики для образного представления состояний энергоблока АЭС [Текст]: А.А. Башлыков // Труды международного форума информатизации. М.: Машиностроение. – 1994. – с. 92-98.

### **Dzhevaga Konstantin Alexandrovich**

The Academy of the Federal Guard Service of the Russian Federation, Russia, Orel  
E-mail: [dka06@mail.ru](mailto:dka06@mail.ru)

### **Mezentsev Alexander Olegovich**

The Academy of the Federal Guard Service of the Russian Federation, Russia, Orel  
E-mail: [aomezentsev@gmail.com](mailto:aomezentsev@gmail.com)

## **Functional model of cognitive-imaginative visualization of heterogeneous data**

**Abstract.** The article discusses the option of solving the problem of analytical processing of large volumes of heterogeneous data source, using the capabilities of a person for the visual analysis of data that are the subject of research in cognitive computer graphics, offered a functional model of the imaging process information. The urgency of the problem consists in the evolution of "open data" concept, the fast growth of the volume of the analyzed data, the need to find regularities obtaining new knowledge, there is no evidence-based methods for the synthesis of adequate source data symbol, taking into account the individual characteristics of perception of the decision maker. The use of human capabilities in the visual analysis of data that are the subject of research in cognitive computer graphics, improves the efficiency of the search patterns of the facts of restrictions volume of source data, a decision on whether to continue the analysis in solving the problem, etc. An approach is proposed to forming a graphic symbol usable for use in systems analyze large volume of data. Describes the factors affecting the process of forming a graphic symbol. Presents a general model a graphic symbol are listed requirements applied to him. Developed on the basis of its functional model of cognitive-imaginative visualization of such data. Developed on the basis of its functional model of cognitive-imaginative visualization of such data. The presence in the functional model reverse communication allows to take into account factors related to the psycho-physiological features the perception of the information and the physical features of the environment of the information display system.

**Keywords:** open data; big data; processing large amounts of data; data mining; computer graphics; cognitive computer graphics; analytical data processing; graphic symbol; information visualization

### **REFERENCES**

1. Barsehyan, A.A. Data Analysis Technology: Datamining, Visialmiming, TextMining, OLAP [Text] / A.A. Barsehyan, M.C. Kypriyanov, V.V. Stepanenko, I.I. Holod. – 2nd ed., revised and enlarged. – SPb.: BHV-Peterburg, 2007. – 384 p.
2. Kim, J.-O. Factor, discriminant and cluster analysis: Translation from English. [Text] / J.-O. Kim, Ch.W. Muller, W.R. Klekka, etc.; Edited by I.S. Enukov. – M.: Finance and Statistics, 1989. – 215 p.
3. Donald E. Knuth The Art of Computer Programming: Volume 3: Sorting and Searching [Text] / Donald E. Knuth – M: Vulyams, 2007. – 824 p.
4. Shashkov V.B. Applied Regression Analysis. Multifactor regression [Text] / V.B. Shashkov – Schoolbook. – Orenburg: GOU VPO OSU, 2003. – 363 p.

5. Afanasev, V.N. Time Series Analysis and Forecasting: schoolbook [Text] / V.N. Afanasev, M.M. Uzbashev. – 2nd ed., revised and enlarged. – M.: Finance and Statistics; INFRA-M, 2010. – 320 p.: img.
6. Tsaplin V.V. multidimensional data visualizing cognitive technologies for decision-making intelligent support [Text]: V.V. Tsaplin, V.V. Vitkovskiy, V.L. Gorokhov // Software and Systems. – 2014. – № 3 (107). – p. 22-25.
7. Manko N.N. Cognitive visualization of pedagogical objects in modern training techniques [Text]: N.N. Manko // Education and Science. – 2009. – №8 (65). – p. 10-30.
8. Gorokhov, V.L. Modern methods of cognitive visualization of multidimensional data [Text] / V.L. Gorokhov, A.A. Lykianez, A.G. Chernov. — Tomsk: Non-commercial development of regional energy fund, 2007. — 216 p.: img.
9. Gorokhov, V.L. Cognitive risk management techniques in the technosphere. Projects and studies. [Text]: V.L. Gorokhov, V.V. Tsaplin, V.V. Vitkovskiy // Bulletin of Civil Engineers: St.P. GASU – 2011 – № 3 – p. 122–127.
10. Tolok, A.V. Graphic images of models in information technology [Text]: A.V. Tolok // Applied Information Science. – 2009. – №4. – p. 31-40.
11. Baburin, V.A. Big Data technologies when new markets, opportunities and challenges [Text]: V.A. Baburin, M.E. Yanenko // Technical and technological service problems. – 2014. – №1 (27) – p. 100-105.
12. Klemenkov P.A. Big data: current approaches to storage and handling [Text]: P.A. Klemenkov, S.D. Kyznetsov. // Proceedings of the Institute for System Program RAN. – 2012. – № 23 – p. 143-156.
13. Big Data Visualization: Turning Big Data into Big Insights. The Rise of Visualization-based Data Discovery Tools. White Paper. Intel IT Center. March 2013.
14. Shneiderman. The big picture for big data: Visualization. Science, 343:730, February 2014.
15. Computer Graphics: methods, models and tools of transformation of graphical information. [Text]: Volume 2. Theoretical basis of image processing and applied aspects of computer graphics: a monograph in 2 volumes. / Ed. Doctor of Technical science A.P. Fisun, Doctor of Technical science I. S. Konstantinov, Doctor of Technical science V.A. Minaev. – Orel: «OSU», «Orel OSU», 2010. P. 238.
16. Bashlukov A.A. Methods of cognitive graphics for figurative representation NPP states [Text]: A.A. Bashlukov // Proceedings of an international forum of information. M.: Mechanical engineering. – 1994. – p. 92-98.