

Когнитивная интерпретация визуализации

А.А. Захарова¹, Е.В. Вехтер², А.В. Шкляр¹
 zaa@tpu.ru|vehter@tpu.ru|shklyarav@tpu.ru

¹ИПМ им. М.В.Келдыша РАН, ²Томский политехнический университет

В работе показаны результаты исследований когнитивной интерпретируемости визуализации и предложены подходы к получению эмпирических оценок практической результативности средств визуализации. Полученные результаты использованы при разработке технологии когнитивной интерпретации гетерогенных данных и необходимых для этого средств визуальной аналитики.

Ключевые слова: визуальная аналитика, визуальная модель, анализ данных, интерпретация, визуальное восприятие.

Cognitive interpretation of visualization

A.A. Zakharova¹, E.V. Vekhter², A.V. Shklyar¹
 zaa@tpu.ru|vehter@tpu.ru|shklyarav@tpu.ru

¹Keldysh Institute of Applied Mathematics RAS, ²Tomsk Polytechnic University

The paper shows some results of studies of cognitive interpretability of visualization and suggests approaches to obtaining empirical estimates of the practical effectiveness of visualization tools. The obtained results were used in developing the technology of cognitive interpretation of heterogeneous data and the necessary tools for visual analytics.

Keywords: visual analytics, visual model, data analysis, visual interpretation, visual perception.

1. Введение

На современном этапе развития средств информационной коммуникации визуализация рассматривается как доминирующая форма информационного взаимодействия между пользователями и источниками информации [12]. Однако, к настоящему моменту по-прежнему не сформирован единый обоснованный подход к изучению и использованию возможностей визуализации в качестве инструмента научного исследования и передачи информации. Одним из препятствий для движения в этом направлении является необходимость проведения объемных междисциплинарных исследований, для которых характерна проблема общего определения объекта исследований и применимости специализированных методов [3].

Специфичным эффектом использования визуализации, согласно [11], является влияние на сознание наблюдателя через формирование новой информационной реальности. Объяснение этого результата основано на изучении влияния иллюзии объективности визуализации на результат ее интерпретации. Для существующих систем визуализации использование указанного эффекта является резервом, привлечение которого затруднено отсутствием систематизированного подхода к визуальной аналитике.

Переход от исключительно иллюстративной визуализации к использованию когнитивного потенциала визуального восприятия происходит медленнее, чем развиваются технологии визуализации. Когнитивная интерпретация визуальных образов предъявляет к создателям подобных образов требования, выполнение которых затруднено необходимостью направленного использования особенностей восприятия и мышления, многие из которых недостаточно изучены. В работе представлены результаты, полученные в результате исследования особенностей визуального восприятия и когнитивной интерпретации абстрактных образов.

2. Комплексность

Одним из перспективных направлений развития теории визуализации следует считать исследование соотношения в интерпретации информативного образа объективной и

субъективной составляющих. В данном случае, к субъективным факторам, влияющим на общий результат когнитивной интерпретации, могут быть отнесены: условия создания визуального образа, специфика восприятия, культурные традиции и другие данные. Следовательно, в визуализации возникает необходимость согласованного использования методов семиотики, лингвистики, психоанализа, феноменологии, герменевтики, системного анализа и т.д.

В исследованиях [1] сформулированы положения, указывающие на необходимость учета при анализе визуальных образов нескольких уровней дополнительной информации: технологического, лингвистического, жанрового, эстетического, идеологического. В работе [13] предложено выделение трех информационных составляющих: функциональной, семиотической и психоэмоциональной, — управление которыми позволяет контролировать результаты когнитивной интерпретации. Выбор составляющих, вносящих вклад в результат процесса когнитивной интерпретации, является основой создания общей схемы визуализации, а также позволяет определить способы ее улучшения.

3. Проблематика визуализации

Исследователями Drexel University's College of Computing & Informatics (CCI) [2] приведен рейтинг нерешенных проблем визуализации. К настоящему времени, этот список может быть уточнен с учетом актуальных задач, к решению которых может быть привлечена визуализация, в т.ч. научная и когнитивная:

1. Юзабилити (Достаточность). Проблема достижения средствами визуального исследования пользовательских характеристик, обеспечивающих возможность получения достоверного решения поставленной задачи исследования лишь при использовании визуальной интерпретации.
2. Перцепция. Задача увеличения информативности и интерпретируемости средств визуальной аналитики в результате направленного использования законов восприятия в схеме визуализации.
3. Предварительная информированность. Проблема избыточного информирования пользователя,

- снижающего когнитивную результативность визуализации.
4. Обучение. Увеличение сложности задач визуального исследования, создаваемых для этого средств визуальной аналитики и необходимость предварительного обучения пользователей.
 5. Качественная мера. Проблема сравнения и выбора средств визуализации из-за отсутствия принятой схемы визуализации и соответствующей ей системы критериев оценки.
 6. Масштабируемость. Задача поиска способов визуализации информации, сохраняющих высокую когнитивную интерпретируемость при любых изменениях детализации или объема данных.
 7. Интерпретируемая эстетика. Исключение или использование субъективного критерия визуальной эстетики для достижения цели исследования.
 8. Динамика. Задача использования восприятия изменяющихся образов для усиления когнитивной интерпретируемости визуализации.
 9. Внутрисистемная логика. Формализация использования визуализации в качестве знаковой языковой системы для получения новых знаний при использовании лишь внутрисистемных операций.
 10. Представление знаний. Решение комплексной задачи эффективного хранения, передачи, получения и использования информации в результате создания ее визуального представления.

4. Методика исследования

Получение ответов на сформулированные вопросы сталкивается с проблемой получения объективных данных, подтверждающих существование процессов, включенных в любую из разработанных схем визуализации [4, 6, 10]. Наличие и корректная интерпретация подобных эмпирических данных создадут условия для построения средств визуализации, обладающих обоснованным и направляемым когнитивным эффектом. Для получения необходимых данных разработана методика исследования возможностей визуального представления, апробированная на нескольких вариантах средств визуального анализа многомерных гетерогенных данных (Рис. 1).

Суть предложенной методики исследования заключается в проведении серий тестовых решений задач визуального анализа, характеризующихся рядом контролируемых ограничений, а также в измерении интервалов времени, соответствующих различным этапам в схеме визуализации. На каждом этапе происходит взаимодействие пользователя с визуальным образом, которое подразумевает любые доступные пользователю

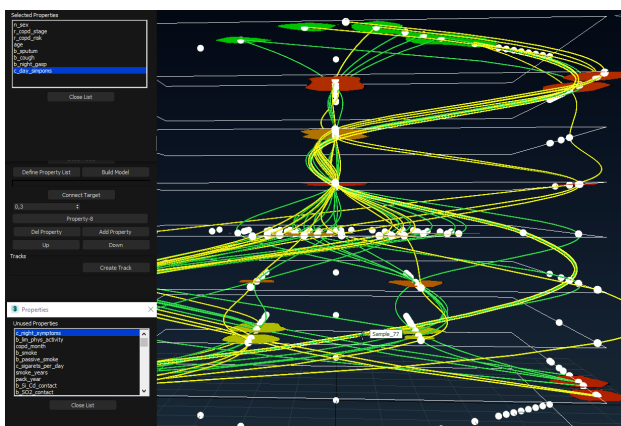


Рис. 1. Применение методики тестирования для одной из визуальных моделей.

операции, за исключением изменения набора анализируемых данных.

Тестовое решение предполагает создание образа исследуемых данных с использованием predetermined способа визуального представления. Выбор первоначального способа визуализации соответствует типу задачи исследования (задача обучения, информирования, принятия решения) и не зависит от предпочтений исследователя. Таким образом, на первом шаге пользователь располагает визуальной моделью данных и знает цель (вопрос) решаемой им задачи [7].

В ходе тестового решения, исследователю, анализирующему образ данных, позволяет формулировать неограниченное число гипотез ответа на вопрос задачи, каждая гипотеза считается шагом анализа. Формулирование верной гипотезы означает завершение тестового решения. Цикличность процедуры решения позволяет автоматически фиксировать продолжительность каждого этапа для последующего анализа результативности визуализации.

Полученные результаты позволили сделать ряд заключений о характере взаимодействия пользователя и визуальной модели в процессе решения задачи анализа. Прежде всего, предложенная методика изучения когнитивных особенностей визуального анализа позволила оценить общее время и влияние ряда факторов на длительность решения. Например, в случае проведения серии измерений с участием одного пользователя это позволило оценить индивидуальные особенности когнитивных процессов. Типичный пример результатов таких измерений (Рис. 2) показывает существенное изменение скорости мышления исследователя в ходе решения однотипных задач.

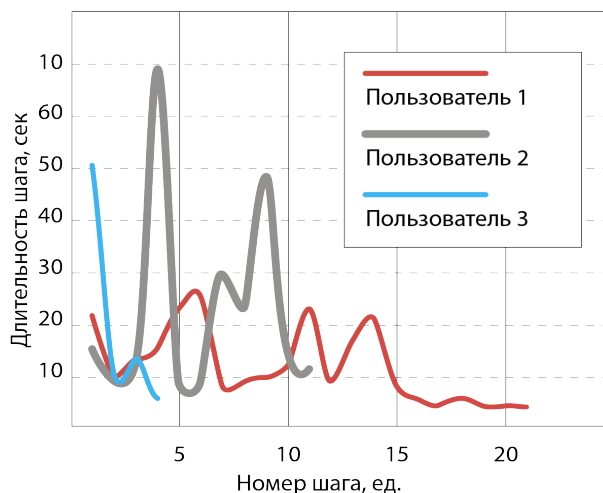


Рис. 2. Длительность шагов в процессе когнитивной интерпретации визуализации.

5. Обучение пользователя

Для получения экспериментальных результатов, позволяющих делать выводы о когнитивных свойствах визуализации, необходима постановка тестовых задач, минимизирующая влияние факторов, относящихся к формальным логическим рассуждениям, опирающимся на существующие у пользователя знания. В качестве тестовых задач пользователю предлагалось исследование наборов многомерных данных, описывающих реальные системы. В соответствии с целью проведения тестовых решений, пользователю было неизвестно происхождение данных и их номинальные значения. Набор визуализированных данных дополнялся искусственно сгенерированными данными,

которые считались аномальным отклонением от реальных значений. Обнаружение такого аномального объекта в наборе исходных данных являлось целью тестового визуального исследования.

Получение с помощью предложенной методики повторяющихся результатов измерений, позволило сделать предположения о некоторых общих закономерностях использования визуализации в качестве инструмента аналитического исследования и получить представление о способах влияния на результаты такого использования. По результатам проведенных измерений, в процессе взаимодействия пользователя и визуальной модели выделены интервалы, на протяжении которых время, затрачиваемое пользователем на построение очередной гипотезы, непрерывно сокращается. Во многих случаях, такие интервалы расположены в начале процесса визуального анализа, однако, могут повторяться и на более поздних шагах решения. Высокая повторяемость полученных результатов и их зависимость от пользователя, принимающего участие в тестовых измерениях, являются основанием для сопоставления с когнитивными аспектами визуализации.

С этой точки зрения, интервалы ускоряющегося формирования новых гипотез, возникающие на начальных этапах схемы визуализации, могут быть интерпретированы как стадии анализа, соответствующие ознакомлению пользователя с инструментом визуального анализа, его особенностями в решаемой задаче, а также накоплению дополнительных сведений, возникающих в результате принятия неверных решений. Следует обратить внимание, на то, что формирование очередной гипотезы ответа, даже в случае ошибочности прежней, многократно повторяется вместе с сокращением затрачиваемого времени. Таким образом, в схеме визуализации, содержащей многократное повторение циклов приема и передачи информации [9], может быть выделена стадия накопления пользователем информации, т.е. этап обучения.

На основании проведенных измерений, показана зависимость продолжительности этапа обучения от способа визуального представления исследуемых данных, а также от собственных характеристик пользователя. Выделение и использование факторов, сокращающих длительность когнитивных процессов, является способом решения многих задач, указанных в качестве проблематики визуализации.

6. Сомнения и рефлексия

Изменение скорости построения новых гипотез, наблюдаемое на протяжении процесса визуального анализа, носит периодический характер. Во время решения тестовых задач пользователь, в рамках проводившегося исследования, не получал сведений из внешних источников, поэтому изменение времени, затрачиваемого на каждом шаге анализа, может быть объяснено изменением соотношения между прямыми и обратными процессами [5]. С точки зрения изучения когнитивных возможностей визуализации, изменяется скорость процедур поиска информации и проверки результатов поиска.

Незначительное снижение скорости решения, с этой точки зрения, свидетельствует об увеличении времени, затрачиваемом на построение и проверку очередной гипотезы, и может интерпретироваться как сомнения в правильности совершаемых действий, вызванное ошибками, допущенными на предыдущих шагах (Рис. 3). Возникновение сомнений приводит к повторному обращению к накопленным знаниям или повторению логических операций с привлечением внутрисистемной логики. Улучшение когнитивной результативности в этом случае может быть достигнуто в результате использования

средств визуализации, предоставляющих оперативный доступ к необходимым сведениям.

В большинстве измерений, резкое снижение скорости формирования пользователем новой гипотезы наблюдалось после серии быстрых ошибочных предположений. Это говорит об отсутствии у пользователя понимания результатов визуальной интерпретации информации, т.е. о противоречивых результатах когнитивных процессов. Значительное снижение скорости анализа возникает с периодичностью, отличающейся у разных пользователей и может быть использовано в качестве одной из характеристик визуализации. Длительность подобного этапа рефлексии, используемого для переосмысления доступной информации и совершенных ошибок, зависит от индивидуальных особенностей исследователя. Особенностью этапа рефлексии является преобладание операций обращения к сведениям, хранящимся в памяти пользователя. Таким образом, сокращение продолжительности этапа рефлексии возможно в случае устранения многократного перекодирования визуальной информации и знаний пользователя в результате выполнения операций поиска и проверки необходимой информации. Одним из возможных решений этой задачи является сохранение в визуальном образе исследуемых данных сведений об ошибочных решениях, сделанных ранее. Подобный прием является способом накопления визуальной информации и его результативность должна исследоваться дополнительно, т.к. не соблюдается граничное требование используемой методики, связанное с неизменностью исходных данных.

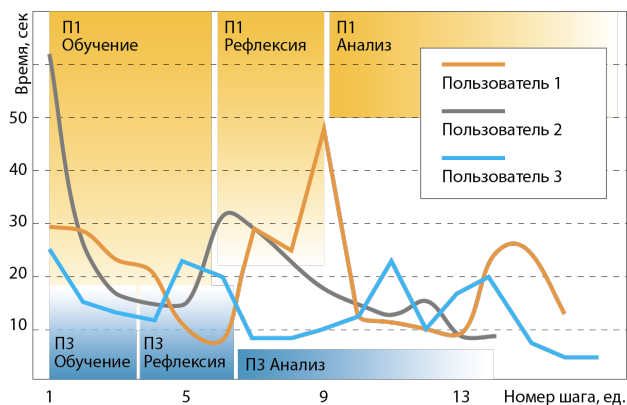


Рис. 3. Стадии обучения и рефлексии.

7. Форма и цвет

На основании измерения времени решения тестовых задач с помощью средств визуального анализа, использующих различные способы визуального представления, проведена оценка некоторых аспектов визуального образа с точки зрения их влияния на скорость формирования новой гипотезы. При проведении серийных измерений предполагалось, что все прочие факторы, кроме исследуемых в соответствующей серии, оказывают на результаты измерения известное и ограниченное влияние.

При сравнении результатов использования визуальных моделей, отличающихся способами цветового кодирования анализируемых данных (Рис. 4), были получены результаты, необходимые для уточнения существующих способов визуализации данных. Полученные результаты интерпретированы как вероятное негативное влияние использования цвета при создании визуального представления данных.

В соответствии с результатами тестовых измерений, цветовая составляющая, во-первых, воспринимается пользователем как дополнительные данные, участвующие в анализе. Это приводит к снижению когнитивной интерпретируемости визуального образа в случае дублирования уже представленных данных или интерпретации пользователем несущественных данных. Во-вторых, пользователь, по результатам опросов участников тестовых решений, во многих случаях отдавал предпочтение информативности формы визуальных объектов, составляющих образ данных, и игнорировал цветовую составляющую. Таким образом, сопоставление результативности использования визуальных моделей в решении различных задач анализа, показало необходимость оправданного использования цветового кодирования при создании визуальных моделей, рассчитанных на когнитивную интерпретацию.

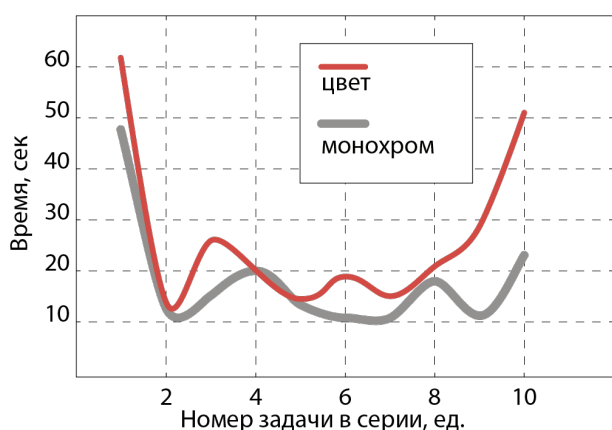


Рис. 4. Влияние цветового кодирования на время решения.

8. Характеристики средств визуализации

На основании разработанной авторами методики проведения визуального исследования данных и созданного обоснования применимости различных средств визуальной аналитики с целью увеличения общей результативности визуального исследования, сделан вывод о необходимости введения собственных характеристик средств визуализации. Предложены определения двух величин, каждая из которых может быть использована для упрощения процессов разработки, выбора и использования средств визуализации.

Для решения вопросов, связанных с выбором средств визуализации, введена комплексная характеристика, объединяющая длительности интервалов времени, затрачиваемых на построение визуальной модели, ознакомление пользователя с ее свойствами и получение ответа на вопрос исследования (Рис. 5). Разработанная методика позволила провести необходимые измерения и получить комплексные характеристики исследованных средств визуализации.

В качестве примера приведены комплексные характеристики трех визуальных моделей, примененных для решения одной задачи анализа данных. Каждый столбец диаграммы (1-3) соответствует одному из тестируемых способов визуализации, комплексная характеристика модели (Рис. 1) соответствует характеристике 2. Характеристика 1 относится к визуальной модели, представляющей облако точек в пространстве с определенной системой шкал, характеристика 3 — относится к визуальной модели, дополненной цветовым кодированием. Нижний сегмент (синий) — этап построения модели, средний (оранжевый) — этап обучения, верхний (серый) — этап решения. Численные характеристики соответствуют средней продолжительности

соответствующих этапов (секунды), полученной с использованием предложенной методики измерений.

Показан вариант нормированных характеристик, позволяющий анализировать когнитивные характеристики визуализации, достигаемые при использовании определенных ресурсов (временных и вычислительных). Одним из практических выводов, которые могут быть сделаны на основании комплексной характеристики, является соотношение между когнитивной интерпретируемостью средств визуализации и их ресурсоемкостью.

Визуальная модель данных является сложным многофакторным объектом и может быть охарактеризована с точки зрения соответствия своему назначению, т.е. процессу когнитивной интерпретации. Для функции визуального представления, формирующей визуальный образ исследуемых данных, предложено введение декомпозиционной оценки восприятия (ДОВ), подтверждающей реализацию в процессе интерпретации образа данных соответствующих принципов восприятия. ДОВ представляет собой многомерный бинарный вектор, содержащий экспертную оценку используемых функций визуального представления принципов восприятия. В Таб.1 приведен вариант подобной оценки для тестируемой модели (Рис. 1).

Таблица 1

Декомпозиционная оценка визуальной модели	
Принцип восприятия	Реализация
Константность	да
Целостность	нет
Обобщенность	нет
Предметность	да
Движение	нет
Цвет	да

Введение этой оценки (ДОВ) для разработанных функций визуального представления позволяет алгоритмизировать процесс решения задач визуального исследования, сократить затраты временных и вычислительных ресурсов и, следовательно, увеличить его результативность.

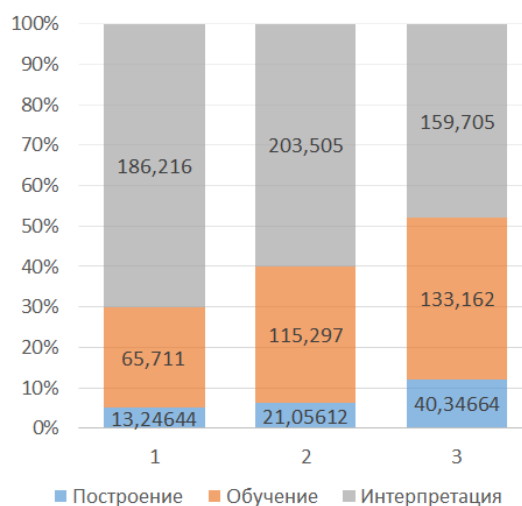


Рис. 5 Примеры комплексных характеристик.

9. Практическое использование

Предложенный подход к исследованию когнитивной интерпретации визуализации позволил разработать средства визуального исследования для ряда практических задач.

Примерами таких решений являются средства визуального анализа гетерогенных данных, предназначенных для оперативного исследования медико-биологических данных [7] (Рис. 6), а также предложенное решение для задачи определения степени изученности заданной предметной области с использованием средств визуализации [8].

В результате применения предложенного подхода к разработке и исследованию возможностей средств визуализации, увеличена результативность решения практических задач. Оценка результативности в этом случае происходила на основании измерения средних интервалов времени, затрачиваемых пользователем на формулирование гипотезы ответа (Таб. 2). Для получения численной оценки результативности предложено сравнение общего времени решения задачи исследования и суммарного интервала обучения и рефлексии. Предполагается, что использование когнитивной интерпретации сокращает время (в проведенных измерениях, на 40-45%), необходимое для общего решения, а завершение цикла обучение/рефлексия соответствует переходу к новой мысленной формулировке цели исследования.

Таблица 2
Оценка результативности решения

Общее время решения, сек	Цикл обучения, сек	Эффект, %
250,9	148,4	40,8
153,2	80,3	47,6
79,5	37,7	52,5
157,6	97,7	38,0
244,9	110,6	54,8
160,4	100,8	37,2

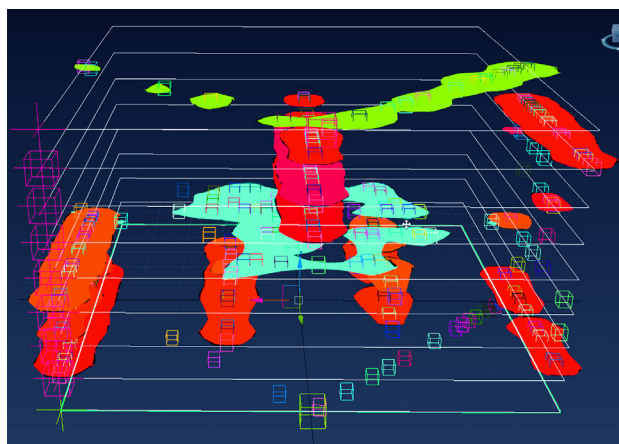


Рис. 6. Модель медико-биологических данных.

10. Заключение

В работе показаны подходы к оценке когнитивной интерпретируемости средств визуализации, используемых при решении практических задач исследования данных. Приведены результаты изучения результативности средств визуализации и показаны некоторые возможные пути управления ее значением. Предложенные характеристики средств визуализации позволяют систематизировать использование визуализации при решении широкого круга задач, имеющих научно-исследовательскую и прикладную направленность. Показана возможность увеличения общей результативности визуального исследования до 35-45% на основании целенаправленного использования возможностей когнитивной интерпретации визуальных образов данных. Перспективным направлением исследований в

рассматриваемой области является согласование когнитивной интерпретируемости визуализации и закономерностей восприятия. Средствами достижения этой цели является определение предложенных характеристик визуализации и факторов, оказывающих на них влияние.

Работа выполнена при поддержке Российского научного фонда, проект 18-11-00215.

11. Литература

- Berger A.A. Seeing is believing : an introduction to visual communication / A.A. Berger, McGraw-Hill Education, 2011. 267 с.
- Chen C. Top 10 unsolved information visualization problems // IEEE Computer Graphics and Applications. 2005. № 4 (25). С. 12–16.
- Elkins J. Interpreting Non-art Images 2001.
- Green T.M., Ribarsky W., Fisher B. Visual analytics for complex concepts using a human cognition model 2008. 91–98 с.
- Pirolli P. Information Foraging Theory: Adaptive Interaction with Information // Information Foraging Theory. 2007. 1–28 с.
- Pirolli P., Card S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis // Proceedings of International Conference on Intelligence Analysis. 2005. (2005). С. 2–4.
- Shklyar A. [и др.]. Visual detection of internal patterns in the empirical data Springer Verlag, 2017. 215–230 с.
- Shklyar A. [и др.]. Visual modeling in an analysis of multidimensional data // Journal of Physics: Conference Series. 2018. № 1 (5). С. 125–128.
- Spivey M. The Continuity of Mind / M. Spivey, 2007. 1-443 с.
- Ward M.O., Theroux K.J. Perceptual Benchmarking for Multivariate Data Visualization 1997. 314–321 с.
- Барт Р. Риторика образа // Избранные работы. Семиотика. Поэтика. / Р. Барт, М.:, 1989. 616 с.
- Розин В.М. Визуальная культура и восприятие. Как человек видит и понимает мир / В.М. Розин, 6-е изд., URSS, 2015. 272 с.
- Шкляр А., Захарова А. Структурный подход к визуализации данных Издательство МАИ-Принт (Москва), 2017. 609–611 с.

Об авторах

Захарова Алена Александровна, Институт прикладной математики им. М.В. Келдыша РАН, zaa@tpu.ru.

Вехтер Евгения Викторовна, Томский политехнический университет, Томск, Россия, vehter@tpu.ru.

Шкляр Алексей Викторович, Институт прикладной математики им. М.В. Келдыша РАН, shklyarav@tpu.ru.