

Визуализация в задачах реконструкции данных

А.В. Шкляр^{1,А}, А.А. Захарова^{2,В}, Е.В. Вехтер^{3,А}

^А Томский политехнический университет, Томск, Россия

^В Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия

¹ ORCID: 0000-0003-4442-7420, shklyarav@tpu.ru

² ORCID: 0000-0003-4221-7710, zaawmail@gmail.com

³ ORCID: 0000-0003-0604-0399, vehter@tpu.ru

Аннотация

Многие прикладные задачи анализа многомерных данных, описывающих состояние реальных физических или иных систем, встречаются с затруднениями, являющимися следствием низкого качества исходных данных, с тем числе пропуски значений, вероятность ошибок или недостоверность измерений. Неполнота данных может стать препятствием для проведения исследования данных с применением многих современных информационных методик. В работе рассматриваются потенциал и возможности инструментов визуальной аналитики для предварительной подготовки, коррекции или полного анализа объемов первичных данных.

Перспективным направлением применения обсуждаемого в работе подхода, характеризуемого целенаправленным привлечением возможностей визуализации, как инструмента анализа данных, и специализированных визуальных метафор является решение задач обработки и интерпретации данных, источниками которых являются киберфизические системы различного уровня сложности, действующие в автономном или частично управляемом режиме. Характерной особенностью таких систем является наличие большого количества датчиков, ответственных за сбор множества видов данных, различающихся как емкостью соответствующих информационных каналов, так и их скоростью и достоверностью. Примерами подобных киберфизических систем являются БПЛА (беспилотный летательный аппарат), роботизированные станции, функционирующие в условиях, препятствующих получению объективного опыта наблюдений (глубоководные роботы), а также множество иных систем мультимодального мониторинга. Эффективное использование данных, собираемых киберфизическими системами мониторинга, является условием для решения большого числа прикладных и исследовательских задач.

Ключевые слова: Визуальная модель, реконструкция данных, метафора, модель данных, интерпретация.

1. Постановка задачи визуального исследования

Множество существующих подходов к изучению и анализу данных, имеющих различное происхождение и сложность, имеют критическую зависимость от качества этих данных (полнота, достоверность, наличие ошибок) [1], [2]. Затруднения, возникающие в случае отсутствия фрагментов объема данных, описывающих, например, динамическую систему, исследуемую пользователем, вносят коррективы в методику исследования. Изменения могут быть направлены на уточнение задачи исследования или же на компенсацию существующих затруднений, основанную на использовании ресурсов, не привлеченных к анализу ранее.

Примерами инструментального подхода, расширяющим возможности анализа данных, являются разнообразные методики визуального исследования, целью которых становится увеличение результативности аналитического процесса в результате эффективного объединения вычислительных, информационных и когнитивных ресурсов, доступных исследователю [3], [4], [5]. Средством исследования в подобных методиках являются визуальные модели данных, которые могут отличаться как вариантами используемых метафор визуализации, так и способами коммуникации между пользователем и исходными данными, реализуемыми с помощью элементов интерфейса визуальной модели. Таким образом, визуальная модель данных выполняет функции интерактивного высокотехнологичного инструмента, применяемого для решения задач анализа данных как в качестве автономного средства, так и в роли компонента информационной системы [6], [7], [8].

Частичное отсутствие значений в исходных данных может быть следствием разнообразных причин и, в общем случае, к отсутствующим данным могут быть отнесены любые значения, если они вызывают сомнения или противоречат другим параметрам. В этой ситуации задача исследования разделяется на два этапа: предварительная экспертиза исходных данных и их последующий анализ, использующий результаты первого этапа. На этом этапе ставятся вопросы оценки значимости отсутствующих данных, реконструкции необходимых информационных элементов, а также отношения к результатам применения дальнейших методик анализа к исследованию объема исходных данных, дополненных результатами процесса восстановления. Объектом интереса являются средства визуализации, позволяющие найти ответы на вопросы обоих этапов или, как минимум, провести экспертную оценку исходных данных.

При построении визуальной модели неполных данных можно опираться на следующую эвристику: значение параметра, принадлежащего одному из объектов в анализируемой выборке, находится в пределах области значений, характеризующим этот параметр, но принадлежащим другим объектам. Иначе говоря, неопределенность отсутствующих значений ограничена (с некоторой точностью) значениями, присутствующими в исходных данных.

Данное предположение накладывает существенные ограничения на круг вопросов, которые могут определять задачу визуального анализа данных. Целью анализа не может быть поиск аномальных значений, экстремальных точек, ошибок и т.п. Наиболее оправданным направлением для усилий разработчиков средств визуализации, следует считать проектирование инструментов, позволяющих аналитику сформировать целостное представление об исследуемом объеме данных. Под представлением, обладающим указанными возможностями, понимается единый объект восприятия – визуальная модель, интерпретация которой указывает аналитику на причины, определяющие появление конкретных значений в исследуемых данных. Например, визуальная модель многомерных данных, использующая идею параллельных координат для представления в едином пространстве исходных данных, описывает актуальное состояние исследуемой системы, параметры которой содержат описания более 300 информативных элементов. Метафора представления позволяет легко выделить объекты с близкими свойствами, при этом и степень «близости», и ее критерии определяются лишь визуально (Рис.1).

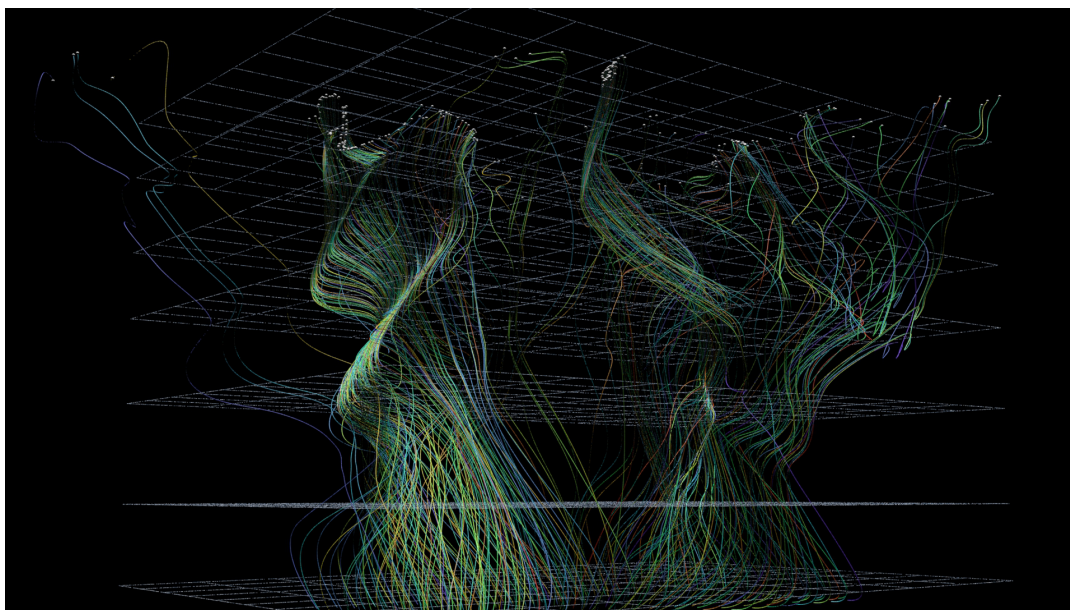


Рис. 1 Визуальная модель многомерных данных, использующая идею параллельных координат

Допустимы некоторые предположения о свойствах такого образа, который выступает в роли модельного объекта, предназначенного для изучения и проведения когнитивного исследования [9], [10]. Наборы экспериментальных данных, характеризующих объекты исследуемой предметной области, могут отличаться своим объемом, условиями получения, состояниями систем объектов или их отдельными свойствами, достоверностью и т.д. При проведении предварительного исследования задачей аналитика может быть получение общих оценочных суждений, не противоречащих как имеющимся данным, так и тем, которые могут появиться позже. Любые избыточные сведения в такой ситуации осложняют деятельность исследователя и требуют к визуальной модели становится упрощение восприятия общей картины данных без потери ее информативности.

Определение элементов исходных данных или их представления, обладающих, в рамках решаемой аналитической задачи, избыточной информативностью, является самостоятельным вопросом, ответ на который становится существенной частью интерпретации исходных данных. В разрабатываемой методике визуального исследования этот этап является обязательным, а его прохождение обеспечивается наличием у инструментов визуального исследования интерактивных возможностей, необходимых пользователю для принятия решения об исключении избыточных данных из поля зрения или, в случае достижения негативных результатов интерпретации, возврата их в работу [11], [12], [13].

2. Метафора различий

При формировании лишь общего представления об исследуемой системе, т.е. без изучения второстепенных данных, из восприятия могут быть исключены такие элементы исходных данных, как идентификаторы событий, значения переменных или повторяющиеся шкалы (Рис.2). Идентификаторы представляют интерес на этапе сопоставления источников данных и особенностей их изменений, но для первоначального анализа эта информация может иметь значение, если известны события, на которых необходимо сконцентрировать внимание. Самостоятельным приемом упрощения является замена плавных сегментов, соединяющих точки данных в описании каждого информативного элемента, на линейные визуальные связи. В этом случае, наглядность представления становится выше, однако, анализ всего объема исходных данных за-

трудняется большим количеством видимых и различимых информирующих образов (точки данных, углы в образах элементов, являющиеся концентраторами внимания).

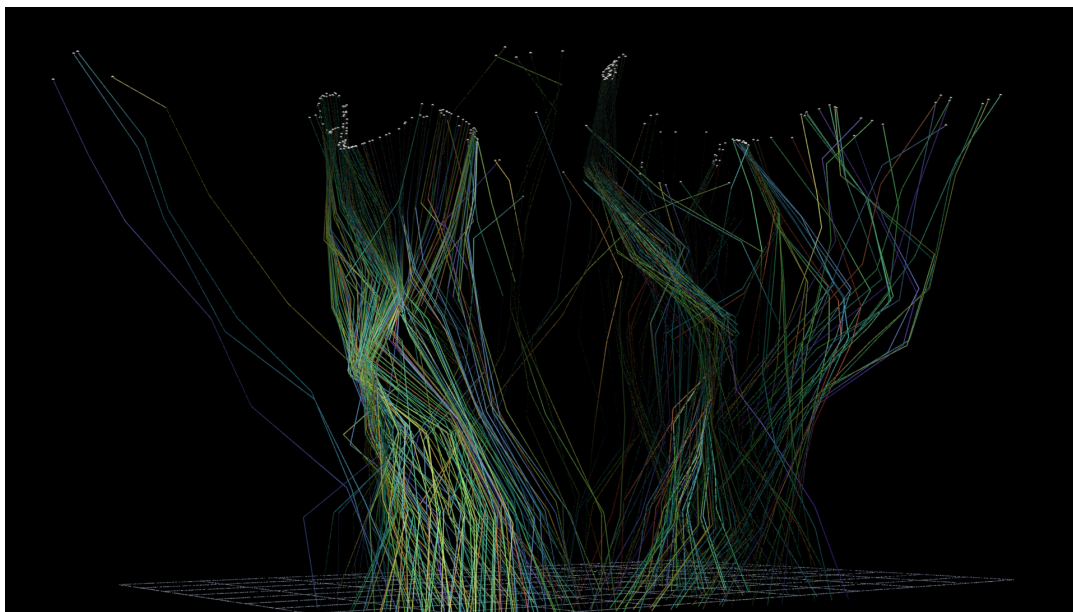


Рис. 2 Вариант упрощения метафоры визуализации

Значения отдельных параметров также могут быть избыточной информацией, т.к. в большинстве случаев основой для получения оценочного суждения становится динамика изменений (или ее отсутствие) параметров, которая может быть оценена при сравнении описаний событий [14], [15]. Следовательно, совместный анализ описаний нескольких событий может быть эффективнее последовательного исследования благодаря лучшей концентрации внимания пользователя. Экспериментальная оценка скорости принятия пользователем решения о значимости отдельных информационных объектов при использовании метафоры параллельных координат подтверждает возможность использования визуальных моделей с упрощенной метафорой визуализации (Рис. 2). Время прохождения этапа экспертной оценки сокращается на 5-15%.

Визуальное представление многомерных данных неизбежно порождает разнообразные когнитивные искажения, связанные с одновременным наблюдением в общем визуальном поле образов данных, масштабирование которых соответствует не значениям их собственных шкал и единиц измерения, а особенностям области представления [16], [17]. Можно предположить, что перспективным выходом станет переход к свободному масштабированию пользователем визуальных образов отдельных измерений. Это позволит исключить (временно или постоянно) из поля зрения шкалы некоторых или всех измерений в многомерных исходных данных, а также изменять их масштабы в соответствии с актуальными потребностями исследователя.

3. Возможности модели отличий

Оптимизация поля зрения исследователя, связанная с исключением второстепенных данных, частично ограничивает разнообразие вопросов, доступных для обсуждения. В частности, анализ значений, присутствующих в исходных данных, становится независимым от абсолютных значений параметров и основывается на относительных величинах. Частным случаем подобной постановки цели этапа экспертизы данных может быть необходимость формирования у исследователя целостного представления о свойствах изучаемых данных [18]. Например, оценка качества данных как соотношения в исходном объеме данных полных и ошибочных описаний объектов (Рис. 3). В этом примере версия метафоры визуализации многомерных данных разработана для предварительной оценки качества исследуемых данных. Каждая линия соответствует обра-

зу многомерного информационного объекта, представленному в «сферическом» варианте параллельных координат. Данные без пропусков образуют линию, полностью расположенную на поверхности «сферы» визуальной модели (цветные линии, каждая линия соответствует одному информационному объекту). Информационные объекты с частично отсутствующими значениями имеют меньшее число точек проецирования на сферическую поверхность модели и поэтому легко выделяются в общем образе данных (голубые линии). Подобная ситуация может быть эффективно использована при наличии возможности автоматического масштабирования визуального представления, однако разрушает понимание диапазона наблюдаемых изменений значений. Важность сохранения такого понимания не является обязательным требованием для инструментов визуализации, т.к. величина диапазона изменений обычно известна исследователю или может быть сообщена ему с использованием минимальных визуальных средств.

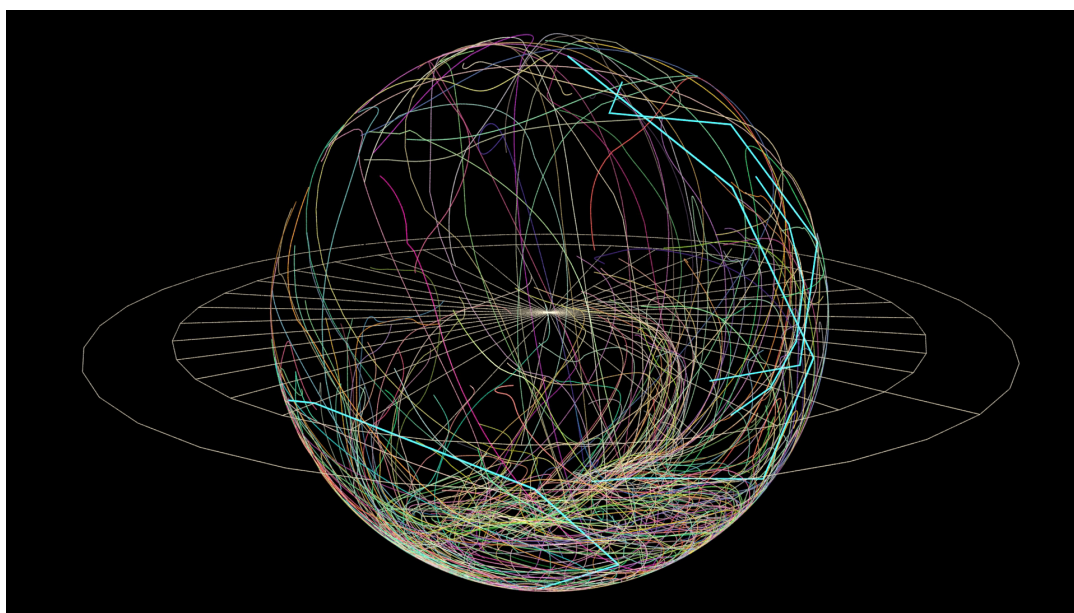


Рис. 3 Метафора визуализации многомерных данных для предварительной оценки качества исследуемых данных

Помимо сравнительного анализа, основанного на сопоставлении наблюдаемых визуальных особенностей исходных данных и сокращающего общее время анализа (до 5 %, в зависимости от самих данных), полезным и неизбежным результатом использования средств визуализации становится прогнозирование значений исследуемых параметров, отсутствующих в исходных данных. Недоступность отдельных значений или значительных фрагментов данных может быть результатом потерь, ошибок измерений или выхода области интереса исследователя за пределы возможностей имеющихся источников данных. В каждой из подобных ситуаций прогнозирование отсутствующих значений опирается на особенности визуального восприятия исследователя, реализованного средствами визуальной аналитики, и, что немаловажно, субъективные знания пользователя, связанные с опытом решения схожих задач анализа.

Реконструкция данных, осуществляемая таким образом, аналогична формулированию гипотез, объясняющих происхождение всего объема данных, и может быть как самостоятельной задачей, так и промежуточным этапом более широкого исследования. Важно, чтобы убедительность визуализации, являющаяся одним из важнейших ее достоинств, не становилась препятствием для критической верификации сформулированных гипотез. Удобным, хотя и не всегда возможным, решением для этого вопроса может стать сравнительный анализ уже не исходных данных, а последствий принятия различных гипотез, визуализированных в качестве альтернативных вариантов в модели данных [11].

4. Метафора отклонений

Углубление в обсуждение тонкостей проектирования инструментов визуализации, связанных с решением специализированных, предметных задач анализа, а также стремление разработчиков визуализации учесть все разнообразие успешных и ошибочных шагов в когнитивном процессе приводят к усложнению процесса визуального восприятия, связанного с необходимостью осмысления насыщенного потока визуальной информации. С другой стороны, обозначенная ранее необходимость исключения избыточной визуальной информации с целью увеличения эффективности решения задач анализа, делают необходимым разработку бесконфликтной системы визуальных выразительных приемов, обладающих достаточной функциональностью в рамках конкретной решаемой задачи анализа. Определение общих рекомендаций, справедливых для некоторого круга задач, может быть основанием для классификации инструментов визуализации, необходимых для предварительного определения целесообразности их применения.

Основанием для оценки и необходимого снижения конфликтности приемов визуализации, применяемых в общем пространстве восприятия может быть выделение независимых, базисных визуальных атрибутов, оказывающих минимальное взаимное влияние. Другими словами, необходимо понимание совместимости информативных потоков в процессе визуального восприятия, безболезненно разделяемых на этапе их интерпретации. Изучение прикладных систем визуализации, в том числе инструментов, применяемых в бизнес-аналитике и научной визуализации, позволяет заметить, что способов отображения данных, совместимых между собой, известно и используется совсем немного [19]. Достаточно легко выделить две большие группы приемов визуализации, которые могут быть использованы как самостоятельно, так и совместно: интерпретация динамики изменений и отклонения от условного равновесия.

Для приемов, основанных на наблюдении за динамикой изменений в исследуемых данных (Рис. 4), характерна высокая концентрация внимания наблюдателя и скорость генерации гипотез. В экспериментальных вариантах визуальной модели, предназначенных для предварительной оценки особенностей исходных данных, использована динамическая метафора, в которой представление образа каждого информативного объекта является последовательным анимированным процессом появления образа. Каждая точка соответствует состоянию одного информационного объекта, присутствующего в исходных данных. Положение точки в пространстве визуальной модели определяется как результирующий вектор в сферическом аналоге параллельных координат (слева). Альтернативный вариант (справа) отличается динамической демонстрацией последовательности состояний объекта, которая существенно превосходит статичную версию по информативности и скорости предварительного анализа. Усложняющийся во времени общий образ исходных данных хорошо интерпретируется при небольшом числе объектов (менее 50), для данных, имеющих невысокую размерность (число параметров в одном описании до 10). В общем случае интерпретация объема исходных данных требует специальной подготовки пользователя.

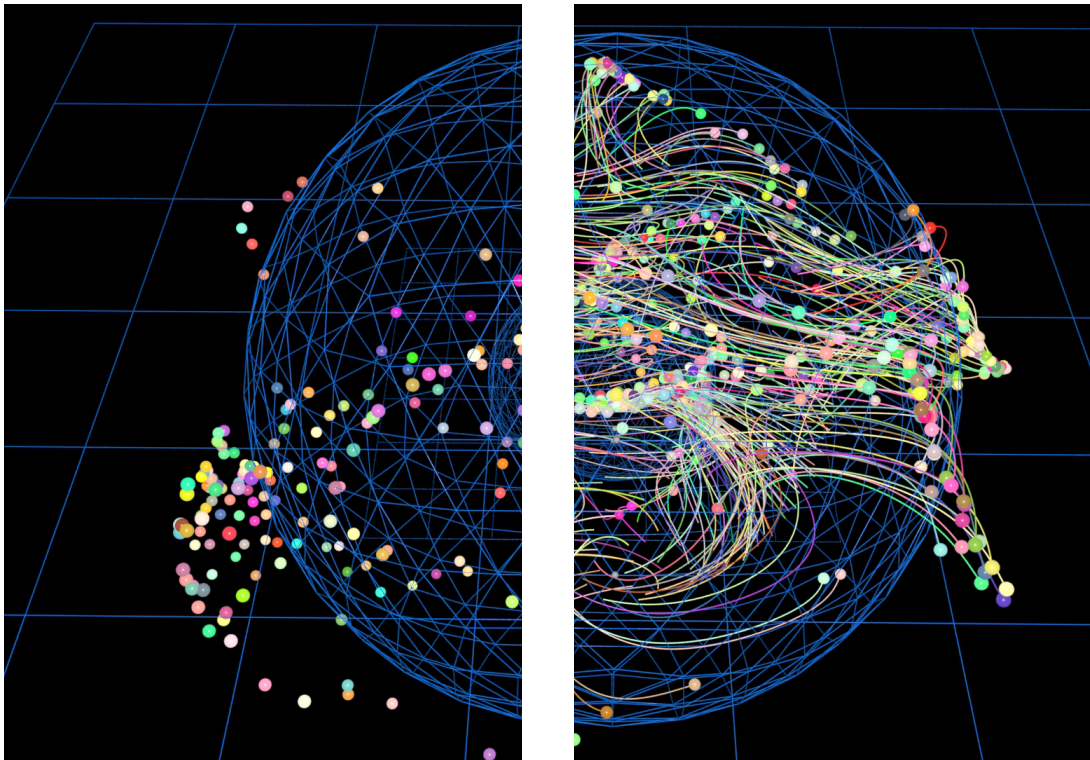


Рис. 4 Экспериментальные варианты визуальной модели для предварительной оценки особенностей исходных данных

Кроме того, наблюдение за изменяющимся образом данных, в котором динамической составляющей может быть не только время, но и любая другая переменная, благодаря специфичным возможностям визуального восприятия инициирует и поддерживает мыслительную деятельность пользователя, связанную с реконструкцией пропущенных или отсутствующих по какой-то иной причине данных. Экспериментальная оценка преимуществ такого подхода осложняется необходимостью определения и систематизации представлений о потенциальном пользователе. В некоторых измерениях, при наблюдении за динамической моделью время предварительного исследования сокращалось на 9-12% в сравнении со статичной визуальной моделью.

Особенностью реконструкции данных такого типа является равноправное использование пользователем для генерации новой гипотезы как результатов актуальной интерпретации, так и данных, оказавшихся в распоряжении исследователя ранее. Указанная особенность конкретизирует область прикладного использования средств визуальной аналитики такого типа и предъявляет требования к знаниям и опыту пользователя. Здесь имеется ввиду, что объем сведений, которыми свободно оперирует человек, работающий с динамической моделью, имеет естественные ограничения, и разработчики метафор визуального представления должны это учитывать.

Идея интерпретации отклонений имеет значительное число версий реализации, среди которых наибольшую популярность имеет оценка пространственных отклонений в образах данных от некоторого условного начального состояния. Другими версиями этой же идеи являются отклонения в цветовом представлении от предварительно определенного нейтрального состояния, а также любое иное отличие экземпляра визуализированных данных от ожидаемого. Разнообразие версий таких отклонений определяет популярность этой группы инструментов визуализации, однако, очень часто приводит к получению метафор визуализации, не учитывающих закономерности визуального восприятия. Примерами таких ошибочных метафор из группы с отклонениями могут быть часто встречающиеся версии цветового кодирования данных, не учитывающие особенности восприятия цвета и яркости, или визуальная несоразмерность мас-

штабов для всех или некоторых параметров в массиве многомерных данных, которая определяет наличие визуальных акцентов.

5. Преимущества метафоры отклонений

Достоинством метафор группы отклонений может считаться получение более точной интерпретации представления данных, сохраняющейся для больших объемов данных, значительного числа анализируемых событий или при наличии существенной хаотичности величин исследуемых параметров, затрудняющей их понимание без привлечения дополнительных приемов визуализации. Это можно объяснить отсутствием ограничений в продолжительности наблюдения за свойствами визуальной модели, а также возможностью одновременного отображения в визуальном поле множества разнообразных данных. Процедура интерпретации такого образа имеет существенные отличия от моделей группы динамических метафор, т.к. становится возможной адаптация скорости когнитивных процессов к индивидуальным потребностям пользователя (Рис. 5). Для оценки преимуществ этого типа визуальных моделей, предложена версия метафоры отклонений, которая предназначена для формирования общего представления об исходных данных, а также для принятия решения о возможности заимствования недостающих значений у информационных объектов, имеющих близкие свойства (желтые и зеленые линии). В предложенной модели отличия свойств информационных объектов визуализируются как отклонение, накапливаемое с удалением от центра модели, поэтому обнаружение схожих объектов не вызывает затруднений.

Потенциальная практическая применимость визуальных моделей, использующих метафору отклонений, выше, чем у динамических или статичных моделей, ориентированных на представление данных. Это связано с эффективным использованием естественных механизмов визуального сопоставления, знакомых большинству пользователей и обладающих высокой скоростью. Сокращение времени, затрачиваемое наблюдателем на выделение объектов повышенного интереса на основании отличий их свойств, от общего объема исходных данных или явно указанных «эталонных» значений, по экспериментальным оценкам, достигает 25-40% и зависит от возможности согласования свойств конкретной метафоры представления и масштаба исследуемых отклонений.

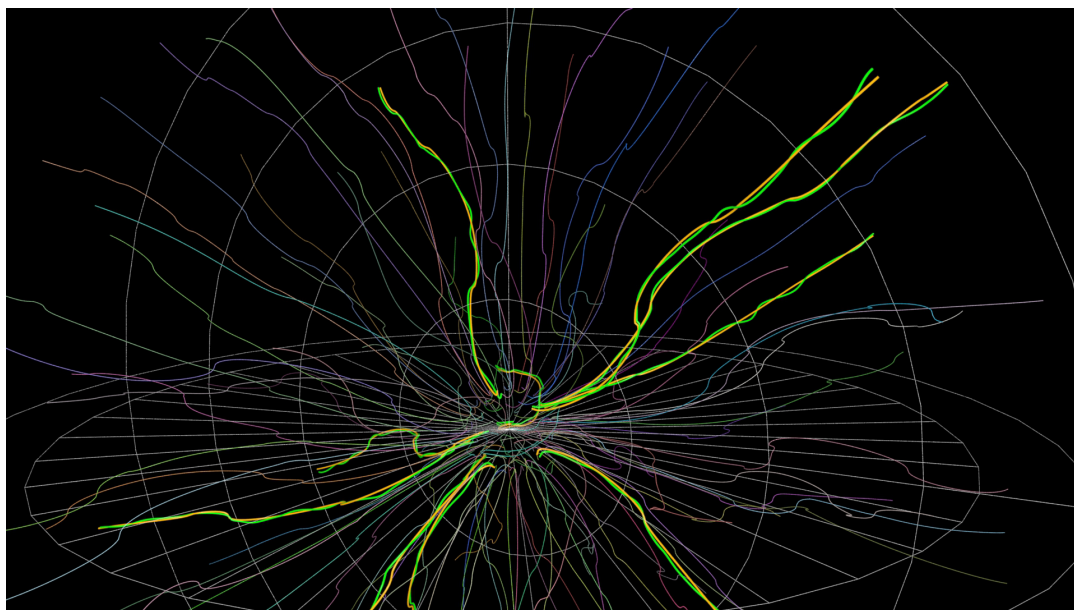


Рис. 5 Вариант визуальной модели, основанной на оценке отклонений

Выделение этих двух групп приемов визуализации данных не делает обязательным выбор в пользу одной из них при разработке инструментов визуальной аналитики. Су-

существует множество прикладных систем визуализации, в которых совмещаются оба подхода: дашборды в BI системах, например, совмещают варианты отображения, акцентированные на возникающих изменениях (мониторы KPI), а также ставшие популярными приемы представления динамических изменений в виде статичных образов (например, широко распространенная метафора «японских свечей» - Candlestick Chart [20]). Представление информации в виде динамически изменяющихся значений KPI ориентировано на анализ не самих значений, а моментов и величины их изменений, т.к. одновременное наблюдение за величиной нескольких переменных создает существенную нагрузку на пользователя [21]. Для ее снижения цифровые значения часто дополняются графическими элементами, указывающими на направленность последних изменений, но это также усложняет общий образ данных. Метафора «японских свечей» задумана как компактная и насыщенная форма представления данных, находящихся в постоянном изменении, и поэтому стала популярным инструментом в решении задач прогнозирования. Недостатком такого подхода следует считать дополнительные требования к подготовке специалистов, использующих эти инструменты, а также ограниченность объема информации, одномоментно передаваемой наблюдателю, которая является следствием перехода к необычной знаковой системе.

6. Роль закономерностей восприятия

Следует отметить, что для представителей каждой из указанных групп приемов визуализации, а также для гибридных вариантов, разрабатываемых из стремления к объединению преимуществ отдельных составляющих, справедливо утверждение о целенаправленном использовании закономерностей визуального восприятия [22]. Понимание роли конкретных особенностей восприятия в поиске ответов на вопросы различного типа позволит вести разработку инструментов визуальной аналитики, основанную на понимании и планировании принципов их действия. К сожалению, примеров ошибочного отношения к применимости средств визуализации в практической области известно очень много и это ставит под сомнение достоинства визуализации в целом.

Для задач реконструкции отсутствующих данных применимость закономерностей визуального восприятия имеет свои особенности, а их игнорирование снижает эффективность инструментов визуальной аналитики. Например, принцип целостности восприятия, который проявляет себя в виде неосознанного стремления наблюдателя к мысленному объединению отдельных визуальных элементов в группу при появлении предположений о наличии общего правила, определяющего их внешний вид. В задаче реконструкции это создает условия для формирования гипотез о сходстве значений отдельных параметров, если общее описание информационных объектов в поле визуальной модели схожи между собой. При отсутствии других оснований для восстановления отсутствующих данных, этот принцип может дать основания для дальнейшего анализа исследуемой системы.

Целостность восприятия может проявлять себя также в виде предположений о существовании информационных структур, объединяющих схожие элементы в визуальном поле и обладающих собственными свойствами, которые также должны быть изучены, но это не всегда справедливо или необходимо (Рис. 6). Например, полезная метафора визуализации, позволяющая одновременное представление как изучаемых данных, так и области допустимых значений, формируемой в результате интерактивного управления моделью, может привести к восприятию и интерпретации визуальной структуры, не предназначенной для этого. Удобным решением для этого затруднения является декомпозиция визуальной модели, позволяющая изучать ее отдельные составляющие. Внесение таких управляющих изменений направлено на получение пользователем дополнительной информации, необходимой для концентрации внимания на специфичных значениях (особенности исходных данных, обнаружение которых при использовании традиционных методик предварительной обработки данных, возможно при усло-

вии наличия гипотезы исследования). Предложенная метафора является инструментом формирования статистических гипотез, проверяемых на следующем этапе анализа данных.

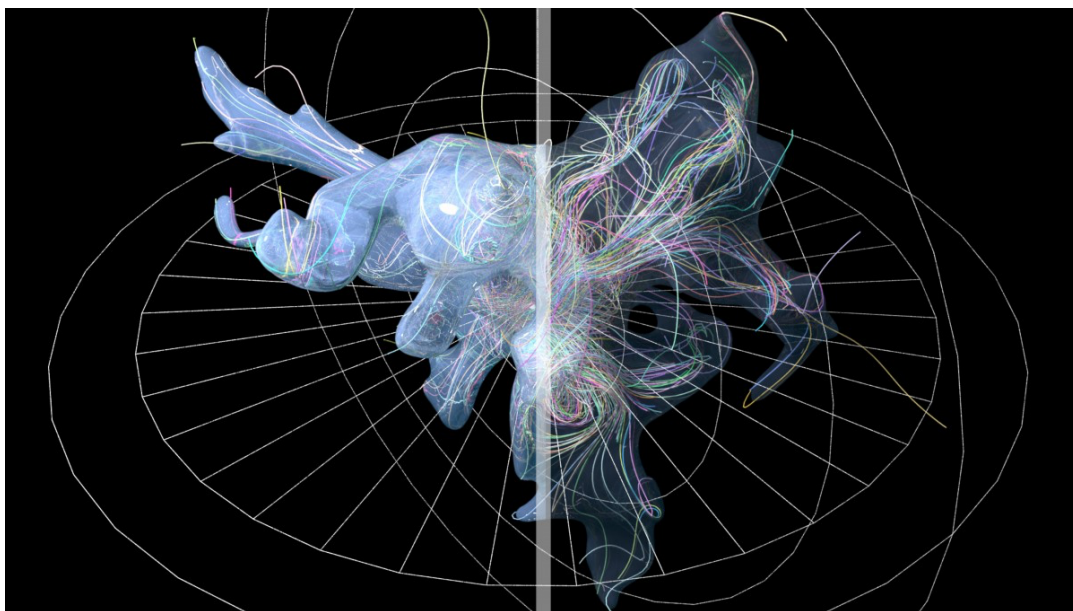


Рис. 6 Визуальная модель многомерных данных, допускающая одновременное представление объема исследуемых данных (анимированные треки – образы информационных объектов в пространстве визуальной модели) и вариации визуализации при внесении любых допустимых изменений в образ данных (синие полупрозрачные области)

Не менее значимая особенность визуального восприятия связана с принципом обобщенности, который во многом определяет процессы распознавания объектов в визуализированной информации или ее отдельных фрагментах. В соответствии с этим принципом, распознавание объекта, известного наблюдателю, происходит даже при наличии искажений в передаваемых данных или иных вариантах информационной зашумленности. Иначе говоря, распознавание объекта заканчивается определением его принадлежности к некоторому классу объектов, принадлежащему системе знаний пользователя. Результатом обобщенного восприятия становится возможность оперирования информацией не об отдельных объектах, а о соответствующих им классах, что, в свою очередь, также позволяет восстанавливать пропуски в исходных данных при объединении признаков объектов на уровне их классификации. В метафоре, использующей идею параллельных координат (Рис. 1), сокращение времени исследования на 5-10% для пользователей знакомых этой идеей, объясняется именно обобщенностью восприятия.

Принцип предметности проявляет себя в создании предпосылок для разделения потока визуальной информации на данные, относящиеся к области актуальной заинтересованности наблюдателя, и элементы, принадлежащие к группе окружения по признакам, установленным самим пользователем. В задачах визуального анализа этот принцип соответствует декомпозиции визуального поля на объекты, чьи параметры и их значения, воспринимаются как самостоятельные информационные структуры, существование которых может быть объяснено в ходе проводимого исследования. Кроме объектов выделяются признаки взаимоотношений выделенных структур с информативным окружением, оказывающим влияние на состояние объектов интереса пользователя. Предметность визуального восприятия является основанием убедительности наблюдаемых образов и, следовательно, данных, использованных для построения изображения. Таким образом, инициируется процесс анализа данных, т.к. предмет-

ность является подтверждением познаваемости и воспроизводимости изучаемых событий.

Кроме того, обратной стороной принципа предметности является возможность обнаружения объектов – информативных элементов визуальной модели, обладающих необычными или непривычными для наблюдателя свойствами (Рис. 3). Это дополняет результаты интерпретации наблюдаемого состояния визуальной модели постановкой новых вопросов и соответствует управлению процессом анализа исходных данных, в том числе его завершением или изменением цели. В обсуждаемых задачах восстановления данных для описания отдельных элементов информационной системы, предметность восприятия может играть значимую роль, т.к. является фактором убеждения пользователя в обоснованности исходных данных и возможности их экстраполяции на описания элементов, в которых отсутствуют необходимые значения. Отсутствие подобной убежденности становится причиной возникновения повторяющихся циклов верификации промежуточных результатов интерпретации, число которых превышает необходимое и зависит от субъективной готовности исследователя к принятию спорных и неочевидных гипотез.

Константность визуального восприятия становится закономерностью, которая имеет важное практическое значение для задач визуализации вообще и для вопросов применения средств визуальной аналитики для восстановления и анализа данных, в частности. Визуальное представление многомерных данных, использующее метафоры сложных геометрических образов в трехмерном пространстве, явным образом использует константность восприятия для предоставления пользователю возможности интерактивного управления пространством модели, в том числе последовательным изменением точек наблюдения за образом данных или продуктивным уточнением метафоры. Принцип константности обеспечивает сохранение пользователем понимания изменяющегося образа модели данных и часто используется для организации возможности произвольного направления анализа многомерных описаний объектов, причем, с учетом субъективного распределения зон интереса в пространстве данных.

При проектировании средств визуализации принцип константности восприятия накладывает ограничения на диапазон допустимых изменений в образе данных. Трансформация воспринимаемого образа данных, любые систематизированные преобразования метафоры представления являются самостоятельным приемом когнитивного исследования, направленным на предоставление пользователю возможности поиска результатов интерпретации, соответствующих поставленным вопросам. Однако, в соответствии с принципом константности, выход за пределы допустимых преобразований приводит к изменению класса, к которому был отнесен объект восприятия или его отдельные части. Это может быть желательным результатом, если изменения в визуальной модели осуществлялись для формирования у наблюдателя качественно новой гипотезы интерпретации, или негативным явлением, если изменение класса информативного объекта стало причиной разрушения системы фактов, связей между ними и соответствующих им знаний, полученных пользователем ранее.

Учет и направленное использование упомянутых выше принципов визуального восприятия, является обязательной стороной проектирования систем визуальной аналитики, так как позволяет сделать действия и умозаключения исследователя систематизированными. Это не означает, что когнитивный процесс будет лишен такой полезной составляющей, как инсайт, но его достижение становится зависящим от накопления пользователем результатов собственной аналитической деятельности. Кроме того, важным этапом когнитивного процесса является его завершение, т.е. состояние, характеризующееся получением ответа на первоначальный вопрос исследования, уверенностью в правильности этого ответа в степени, достаточной для перехода к новому вопросу и новой цели исследования. Время достижения пользователем такого состояния определяет эффективность используемых им инструментов и соответствующих им методик

анализа и определяется не только свойствами решаемой задачи и средств визуальной аналитики, но и взаимодействием этих средств с пользователем.

Обсуждение особенностей взаимодействия визуализированной информации с восприятием и мышлением исследователя, с учетом современных возможностей компьютерной графики и информационных технологий в целом, не может остановиться на обсуждении применимости классических закономерностей восприятия в вопросах интерпретации визуальных моделей данных. Это связано с существенным увеличением разнообразия выразительных приемов, которые могут быть использованы в задачах визуализации, а также появлением специфичных инструментов визуализации, относящихся исключительно к цифровым технологиям и не имеющих близких аналогов в реальности [23].

7. Интерпретация движения

Примером такого направления в визуализации является применение анимации - движения элементов, составляющих композицию образа данных, или любых других преобразований метафоры, или изменений в деятельности способов представления, вошедших в общую метафору, синхронизированных в восприятии пользователя со временем наблюдения (Рис. 4). Если оценивать перспективность привлечения анимации к задачам визуальной аналитики, то к ней во многом применимы суждения, связанные с действием закономерностей восприятия, однако существует целый ряд специфичных проявлений влияния анимации на интерпретацию визуализированной информации, которые могут как усилить, так серьезно уменьшить активность когнитивного процесса. Проблемность анимации связана с активным привлечением субъективного понимания движения, которое имеет противоречивые свойства, заметно усиливающиеся, если разрабатывается инструмент визуальной аналитики для коллективного использования.

Одной из значимых закономерностей, которые используются создателями анимированных изображений самого разного типа, является подготовленность восприятия к следующему движению или, иначе говоря, ожидание, сформированное в результате осмысления уже произошедших событий. В кинематографе этот факт используется для сокращения времени, необходимого для перехода очередному событию, но в визуальной аналитике подготовленность восприятия является индикатором, соответствующим пониманию зрителем причинно-следственных связей, влияющих на происхождение изучаемых данных. Это понимание может быть связано со знаниями, имевшимися у пользователя ранее, или новой гипотезе, верификация которой происходит при помощи визуальной модели.

Прямым следствием активной роли информированного ожидания в интерпретации визуальной модели данных является быстрое обнаружение несоответствий между предполагаемым развитием событий в визуальном поле и действительным поведением изучаемой системы (Рис. 4). Нарушенное ожидание является эффективным способом обнаружения ошибок в двух направлениях: в исходных данных, т.е. позволяют оценивать качество исследуемых данных, или в знаниях пользователя. Кроме того, сформированное у пользователя ожидание в интерпретируемом движении является индикатором достижения степени понимания, осознаваемой как гипотеза происхождения данных. Эффективность, в данном случае, объясняется отсутствием необходимости в дополнительных усилиях со стороны разработчиков инструментов визуализации для привлечения внимания к указанным ошибкам, т.к. переключение внимания на анализ ошибок происходит у пользователя естественным образом.

Дополнительным достоинством анимации является ее активная роль в поддержании высокого уровня вовлеченности наблюдателя в исследование. В данном случае, говорится о влиянии психоэмоционального состояния пользователя на процесс решения задачи визуального анализа. Иначе говоря, если длительная концентрация внимания

пользователя на работе приводит к замедлению или остановке когнитивного процесса, то анимация, как часть метафоры визуализации, способна регулировать вовлеченность пользователя за счет таких приемов как ритмичность восприятия, переключение между режимами наблюдения и интерпретации, а также привлечение персональной заинтересованности [24].

8. Методологический подход к задаче визуального исследования

В общем случае, методологический подход к задаче исследования неполных данных состоит в выборе одного из возможных алгоритмов решения, подготовке условий для его реализации и разработки инструментария – соответствующих задаче средств визуальной аналитики и качественной оценки результата. В качестве двух возможных алгоритмов проведения исследования неполных данных могут быть указаны следующие альтернативные друг другу варианты действий:

А. Исключение из процесса исследования данных, которые недоступны для изучения традиционными или иными путями из-за отсутствия значений в части образцов. Построение модели данных с учетом предположения о несущественной роли отсутствующих данных в процессе поиска ответов на вопросы, стоящие перед пользователем. Важно, что в этом случае визуализация должна использоваться, в том числе, и для формирования у исследователя уверенности в возможности продолжения решения задачи анализа неполного набора данных. Отсутствие убежденности в правильности действий может быть компенсировано на этапе оценки получаемых результатов анализа. Однако, переходу к верификации результатов предшествует этап генерации гипотез, при прохождении которого наличие субъективной неуверенности у пользователя приводит к избыточному многократному выдвижению новых гипотез, что существенно замедляет решение задачи анализа. Следует отметить, что потенциальная применимость указанного алгоритма решения во многом опирается на предметность и обобщенность визуального восприятия.

В. Замена отсутствующих значений в описаниях информативных элементов данных на предполагаемые, т.е. гипотетически возможные. Частным случаем такого варианта решения является введение на соответствующих шкалах специальной точки «данные отсутствуют» и проведение анализа с учетом возможности состояния источников данных, когда они не генерируют необходимые значения. Основанием для реконструкции в этом варианте алгоритма решения является гипотеза о близости отдельных значений многомерных данных, при условии достаточной схожести значений в большинстве других имеющихся измерений. Гипотеза неочевидная и может быть обоснована при решении конкретной задачи визуальной оценкой сходства полных наборов значений для имеющихся в распоряжении пользователя информативных элементов, при условии их достаточного количества в исходном наборе данных. В этом случае визуализация также становится удобным средством обоснования применимости выбранного подхода реконструирования при быстром предварительном анализе благодаря привлечению таких закономерностей визуального восприятия как целостность и константность.

Разработка инструментов визуализации, с учетом этих предположений, направлена, в первую очередь, на определение эффективной метафоры визуального представления, для которой будет характерным наличие возможности одновременного отображения в визуальном поле полного объема имеющихся данных для формирования у наблюдателя целостного представления о нем и его особенностях. При этом, в обоих алгоритмах решения задачи анализа разрабатываемые метафоры действуют схожим образом, т.е. дают оценку удаленности (видимое различие) или близости (сходство образов) информативных элементов в многомерном пространстве данных, но имеют противоположные цели.

А. Для алгоритма исключения необходима метафора представления, делающая акцент на оценке разнообразия. Т.е. в случае несущественного влияния разнообразия неполных данных на их смысл возможно принятие решения о допустимости их исключения из общего объема исходных данных.

В. Для алгоритма замены метафора визуализации должна способствовать эффективному поиску объектов с близкими свойствами, чтобы дать основания для заимствования отсутствующих данных.

Необходимо отметить, что возможна ситуация, в которой для обоих вариантов будет применима одна и та же метафора визуализации, т.к. требования к инструментам анализа схожи. В этом случае, корректность и результативность их применения будет зависеть от подготовленности пользователя, его понимания цели визуального анализа и субъективной эффективности используемых выразительных средств.

На этапе качественной оценки правильности выбора и успешности применения средств визуализации для решения любой задачи анализа данных появляется необходимость определения успешности преодоления множества объективных проблем и ошибок, допущенных пользователями или разработчиками. Общими проблемами для самых разных данных и задач всегда становятся следующие:

А. Представление нечисловых данных. Традиционный прием использования категориальных шкал измерения может создавать определенные трудности в ситуации, когда для визуализации информационного элемента исходных данных требуется несколько различных шкал такого типа. Технически, никаких трудностей нет, и формальная процедура построения образа не встречает препятствий. Однако, для пользователя, перед которым стоит цель визуального определения степени сходства или различия между элементами данных, этот вопрос станет достаточно сложным, т.к. понимание расстояния в смысловом пространстве, определяемом категориальными шкалами, редко соответствует субъективному опыту пользователя. В этом случае целостность образа информативного элемента, сформированная в восприятии наблюдателя, может стать основанием для отнесения его к определенному классу объектов, а использование категориальных шкал может быть заменено присвоением объектов воспринимаемых атрибутов.

В. Оценка значимости. Соотношение в визуальном представлении значимости отдельных свойств является одновременно резервом любой метафоры представления и ее слабой стороной. В данном случае, затруднения вызывает предварительное определение факторов, управляющих вниманием пользователя. На практике это соответствует высокой вероятности того, что от внимания наблюдателя ускользнут значимые элементы образа данных, или, наоборот, незначительные колебания значений станут причиной глубокого когнитивного процесса и исказят общий смысл изучаемых данных.

С. Абстрактность или сложность метафоры. Геометрические свойства визуального образа представляют собой модельное представление исходной информации и крайне редко имеют непосредственное отношение к реальным свойствам данных. Метафора представления для любых многомерных данных в восприятии обычного наблюдателя никак не соотносится с восприятием самих данных. Простым наглядным примером такого положения дел может быть обычная географическая карта, на которой информация о высоте точки поверхности представлена в виде цветового кодирования. Для более сложных ситуаций, когда изучаемые данные организованы сложнее или имеют неизвестные пользователю свойства, любая метафора порождает образ, обладающий высокой визуальной сложностью и степенью абстрактности. Этот факт, в общем случае, не является проблемой, однако, с учетом закономерностей восприятия, могут появиться неожиданные результаты при определении наблюдателем категории объекта или его информирующих свойств.

Прямым следствием для указанных сложностей является возрастающая значимость опыта, восприятия и когнитивных способностей конкретного исследователя. Возникает

обязательный этап определения особенностей информационной коммуникации, осуществляемой с использованием визуальных моделей. Выбор способа представления данных может быть сделан управляемым и контролируемым, чтобы делегировать пользователю возможность интерактивной и субъективной настройки модели под свои потребности или особенности восприятия.

9. Применение визуальных метафор для обработки и интерпретации данных киберфизических систем

Перспективным направлением применения обсуждаемого в работе подхода является решение задач обработки и интерпретации данных, источниками которых являются киберфизические системы (КФС) различного уровня сложности, действующие в автономном или частично управляемом режиме. Особенностью таких систем является наличие большого количества датчиков, ответственных за сбор множества видов данных, различающихся как емкостью соответствующих информационных каналов, так и их скоростью и достоверностью. Примерами подобных киберфизических систем являются БПЛА, роботизированные станции, функционирующие в условиях, препятствующих получению объективного опыта наблюдений (глубоководные роботы), а также множество иных систем мультимодального мониторинга.

Для всех этих задач характерна необходимость получения результатов анализа при условии невозможности проведения повторных измерений, или при наличии высокой зашумленности, или потерь фрагментов исходных данных. При оценке преимуществ рассматриваемого в работе подхода к решению задач визуального исследования, был разработан набор визуальных метафор, предназначенных для оценки качества и предварительного анализа разнородных данных, источником которых являются системы мониторинга. Тестовые данные имели различный объем и уровень сложности: от 20 до 300 записей, в каждой из которых могло содержаться 5-18 параметров различного типа, к которым относятся измерения температуры, давления, газосодержания и т. п. (Рис. 7). Использование унифицированных метафор применительно к анализу данных, полученных КФС при мониторинге окружающей среды, позволило сократить время проектирования средств визуального исследования. Например, при увеличении объема исходных данных (до 500 записей) время построения визуальной модели не изменяется. Кроме того, подобная унификация обеспечила снижение затрат на подготовку визуального аналитика к работе над задачами в незнакомых направлениях.

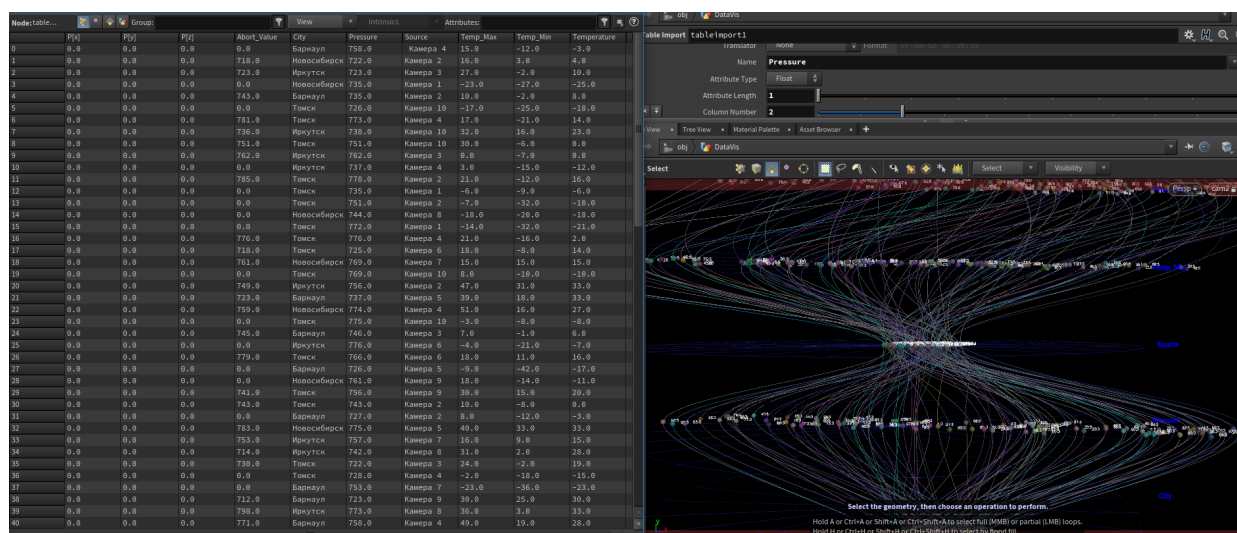


Рис. 7 Пример тестовых данных и выбранной для их анализа визуальной метафоры

Использование метафоры параллельных координат для рассматриваемой задачи обусловлено необходимостью привлечения опыта решения аналогичных задач, полу-

ченного пользователями ранее. Кроме того, эффективность предложенного средства визуализации, основывается на использовании метафоры отклонений, показывающей хорошие результаты при анализе многомерных данных, схожих по параметрам с тестовыми. В качестве приема визуализации, необходимого для указания на отсутствующие значения, используется исключение из поля зрения пользователя соответствующих узлов модели, а также изменение яркости фрагментов образов информативных элементов с недостающими значениями (Рис. 8). При этом, одновременное присутствие в поле зрения имени узла и затененного фрагмента образа элемента используется как активный указатель на область модели, требующей внимания пользователя.

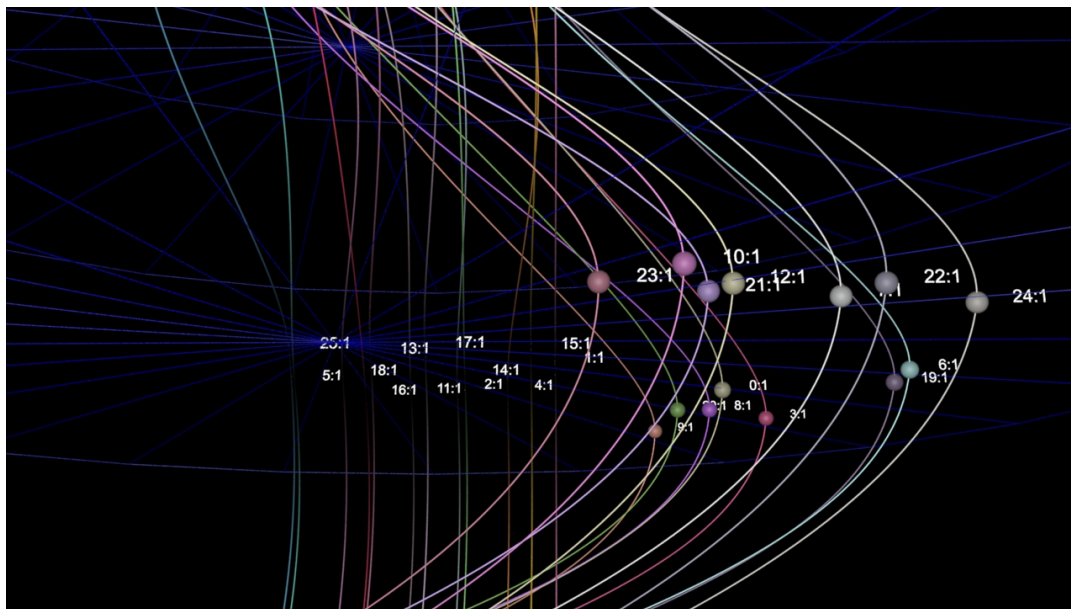


Рис. 8 Визуализация отсутствующих значений. Имя каждого узла модели содержит номер информативного элемента, а также индекс значения в записи

В некоторых ситуациях, например, при высокой зашумленности исходных данных или значительном объеме пропущенных значений, характерной для анализа данных, собираемых КФС, приема цветового кодирования становится недостаточно, т.к. внимание пользователя рассеивается в процессе интерпретации неоднородного насыщенного образа. Для снижения нагрузки на восприятие пользователя использован прием, снижающий детализацию визуальной модели без уменьшения ее информативности (Рис. 9). В предложенном решении, узлы визуальной модели, не имеющие значений, смещаются в специальную область образа (здесь, центральная точка). Для пользователя указателем на область внимания становятся не узлы, а пересечения образов полных и поврежденных элементов.

Для решения задачи восстановления пропущенных значений (когда в рамках серии экспериментов нет полной идентичности набора измерительных средств КФС или возникают технические и иные сбои при записи данных) при использовании рассматриваемой метафоры (Рис.10) необходимо выбрать образы полных информативных элементов, близкие по значениям большинства параметров к образам элементов с пропущенными данными. Для этого достаточно выделить поврежденный элемент и определить номер информативного элемента, пригодного для построения гипотезы заимствования. Для упрощения процесса построения гипотезы заимствования параметр с поврежденными данными смещен в нижнюю часть модели. Это позволяет разделить полные и поврежденные элементы, сохраняя общее представление о всех доступных для анализа значениях. Наибольшую пользу предлагаемая методика приносит в ситуациях, когда источника прямолинейного заимствования нет в исходных данных и для формирования необходимой гипотезы необходимо полное представление об исходных дан-

Заключение

В работе проведен анализ многочисленных факторов, оказывающих существенное влияние на эффективность применения средств визуальной аналитики при решении прикладных задач анализа многомерных данных различного типа. В качестве тестового примера такой задачи рассматривалась проблема подготовки и исследования многомерных данных, собираемых киберфизическими системами различного типа. Особенностью таких данных является отсутствие части значений в описаниях некоторых информативных объектов. Предложен методологический подход, который опирается на преимущества визуального восприятия при генерации гипотез интерпретации и последующего анализа таких данных.

Показана возможность проектирования средств визуализации, применимых для решения прикладных задач, с возможностью интерактивного определения наиболее перспективного алгоритма решения задачи исследования данных указанного типа, а также для других прикладных задач, в которых достижение цели исследования осложнено отсутствием апробированных методик привлечения потенциала современных информационных технологий. Показана роль визуализации и закономерностей визуального восприятия как в процессе исследования исходных данных, так и в контроле над эффективностью использования субъективного потенциала пользователя средств визуальной аналитики.

В задаче реконструкции данных, собираемых КФС, метафора отклонений может быть использована на этапе оценки качества данных в случае многократного применения инструментов визуальной аналитики. Подобная ситуация характерна для анализа данных, получаемых от систем мониторинга, в т.ч. с большим числом разнотипных датчиков. Апробация разрабатываемой методики исследования, показала, что сформированное в этом случае понимание пользователем ожидаемых значений делает модель отклонений инструментом, увеличивающим скорость решения задачи анализа (от 25% по результатам экспериментальных оценок). Метафора отличий показывает хорошие результаты при оценке возможности заимствования значений у сходных информационных объектов и потому также рекомендуется в качестве основы для проектирования средств визуализации на этапе предварительного исследования. Кроме того, версии визуальных моделей, опирающиеся на обобщенность восприятия и интерпретацию движения, позволяют повысить скорость решения задачи исследования до 40%.

Работа выполнена при поддержке Российского научного фонда, грант № 23-19-00342, <https://rscf.ru/en/project/23-19-00342/>

Список литературы

1. А. В. Шкляр, А. А. Захарова, и Е. В. Вехтер, «Методика решения задач анализа данных при использовании аналитических визуальных моделей», Научная визуализация, т. 9, вып. 4, сс. 78–88, 2017.
2. C. Chen, «Science Mapping: A Systematic Review of the Literature», J. Data Inf. Sci., т. 2, вып. 2, сс. 1–40, мар. 2017, doi: 10.1515/jdis-2017-0006.
3. A. V. Shklyar, A. A. Zakharova, E. V. Vekhter, и A. J. Pak, «Visual modeling in an analysis of multidimensional data», J. Phys. Conf. Ser., т. 944, с. 012127, янв. 2018, doi: 10.1088/1742-6596/944/1/012127.
4. J. F. Rodrigues Jr., L. A. S. Romani, A. J. M. Traina, и C. Traina Jr., «Combining Visual Analytics and Content Based Data Retrieval Technology for Efficient Data Analysis», в 2010 14th International Conference Information Visualisation, IEEE, июл. 2010, сс. 61–67. doi: 10.1109/IV.2010.101.
5. A.E. Bondarev, V.A. Galaktionov. Generalized Computational Experiment and Visual Analysis of Multidimensional Data (2019). Scientific Visualization 11.4: 102 - 114, DOI: 10.26583/sv.11.4.09.
6. C. Vieira, P. Parsons, и V. Byrd, «Visual learning analytics of educational data: A systematic literature review and research agenda», Comput. Educ., 2018, doi: 10.1016/j.compedu.2018.03.018.
7. H. Chen и др., «Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration», IEEE Trans. Vis. Comput. Graph., т. 21, вып. 9, сс. 1072–1086, сен. 2015, doi: 10.1109/TVCG.2015.2410278.
8. V. L. Averbukh и др., «Virtual reality as an instrument of computer visualization», в 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), окт. 2019, сс. 0786–0792. doi: 10.1109/SIBIRCON48586.2019.8957854.

9. М. А. Yalçın, N. Elmqvist, и В. В. Bederson, «Cognitive Stages in Visual Data Exploration», в Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16, 2016. doi: 10.1145/2993901.2993902.
10. Касьянов В.Н., «Методы и средства визуализации информации на основе атрибутированных иерархических графов с портами», Сибирский Аэрокосмический Журнал, т. 24, вып. 1, Art. вып. 1, 2023.
11. A. Batch и N. Elmqvist, «The Interactive Visualization Gap in Initial Exploratory Data Analysis.», IEEE Trans. Vis. Comput. Graph., т. 24, вып. 1, сс. 278–287, янв. 2018, doi: 10.1109/TVCG.2017.2743990.
12. H. Guo, S. R. Gomez, C. Ziemkiewicz, и D. H. Laidlaw, «A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights.», IEEE Trans. Vis. Comput. Graph., т. 22, вып. 1, сс. 51–60, янв. 2016, doi: 10.1109/TVCG.2015.2467613.
13. R. Pienta и др., «VIGOR: Interactive Visual Exploration of Graph Query Results.», IEEE Trans. Vis. Comput. Graph., т. 24, вып. 1, сс. 215–225, янв. 2018, doi: 10.1109/TVCG.2017.2744898.
14. R. J. Crouser, L. Franklin, A. Endert, и K. Cook, «Toward Theoretical Techniques for Measuring the Use of Human Effort in Visual Analytic Systems», IEEE Trans. Vis. Comput. Graph., 2017, doi: 10.1109/TVCG.2016.2598460.
15. I. R. Akhmadeeva, Y. A. Zagorulko, и D. I. Mouromtsev, «Ontology-Based Information Extraction for Populating the Intelligent Scientific Internet Resources», в Knowledge Engineering and Semantic Web, А.-С. Ngonga Ngomo и Р. Кřemen, Ред., Cham: Springer International Publishing, 2016, сс. 119–128.
16. A. V. Shklyar, A. A. Zakharova, E. V. Vekhter, и D. A. Zavyalov, «Visual detection of internal patterns in the empirical data», в Communications in Computer and Information Science, Volgograd, Russia: Springer Verlag, 2017, сс. 215–230. doi: 10.1007/978-3-319-65551-2_16.
17. C. Chen и M. Song, «Visualizing a field of research: A methodology of systematic scientometric reviews», PloS One, т. 14, вып. 10, с. e0223994, 2019.
18. C. Ware, Foundation for a Science of Data Visualization. 2012. doi: 10.1016/B978-0-12-381464-7.00001-6.
19. I. Meirelles, «Diagramming As A Strategy For Solving Graphic Design Problems», Просмотрено: 29 февраль 2020 г. [Онлайн]. Доступно на: https://www.academia.edu/2749924/Diagramming_As_A_Strategy_For_Solving_Graphic_Design_Problems
20. С.Нисон, Японские свечи. Графический анализ финансовых рынков. Альпина Паблишер, 2017.
21. J. H. Larkin и H. A. Simon, «Why a Diagram is (Sometimes) Worth Ten Thousand Words», Cogn. Sci., т. 11, вып. 1, сс. 65–100, 1987, doi: 10.1111/j.1551-6708.1987.tb00863.x.
22. E. Branchini, U. Savardi, и I. Bianchi, «Productive thinking: The role of perception and perceiving opposition.», Gestalt Theory, т. 37, вып. 1, сс. 7–24, 2015.
23. A. Dasgupta и др., «Familiarity Vs Trust: A Comparative Study of Domain Scientists' Trust in Visual Analytics and Conventional Analysis Methods.», IEEE Trans. Vis. Comput. Graph., т. 23, вып. 1, сс. 271–280, янв. 2017, doi: 10.1109/TVCG.2016.2598544.
24. J. Choi, S. Jung, D. G. Park, J. Choo, и N. Elmqvist, «Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization», Comput. Graph. Forum, т. 38, вып. 3, сс. 249–260, июн. 2019, doi: 10.1111/cgf.13686.

Visualization in Data Reconstruction Tasks

A.V. Shklyar^{1,A}, A.A. Zakharova^{2,B}, E.V. Vekhter^{3,A}

^A Institute of Control Sciences of Russian Academy of Sciences

^B V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

¹ ORCID: 0000-0003-4442-7420, shklyarav@tpu.ru

² ORCID: 0000-0003-4221-7710, zaawmail@gmail.com

³ ORCID: 0000-0003-0604-0399, vekhter@tpu.ru

Abstract

Many application tasks of multidimensional data analysis which describe the state of real physical or other systems face with difficulties. This is a consequence of the low-quality source data, including missing values, the probability of errors or unreliability of measurements. Incomplete data can become an obstacle for research using many modern informational methods. The current work examines the potential and capabilities of visual analytics tools for preliminary preparation, correction or complete analysis of primary data volumes.

A promising area of application of the approach discussed in the study is the targeted use of visualization capabilities as a data analysis tool. The implementation of specialized visual metaphors is used to solve problems of processing and interpreting data, the sources of which are cyberphysical systems of different complexity levels. Such systems operate in an autonomous or partially controlled mode. A characteristic feature of these systems is the presence of a large number of sensors that collect various types of data. Such data differ in the capacity of the corresponding information channels, their speed and reliability. Examples of such cyberphysical systems are unmanned aerial vehicles (UAVs), robotic stations, and multimodal monitoring systems. These systems can function in conditions where it is difficult to obtain objective observation experience (deep-sea robots). The effective use of data collected by cyberphysical monitoring systems is a condition for solving a large number of application and research tasks.

Keywords: visual model, data reconstruction, metaphor, data model, interpretation.

References

1. V. Shklyar, A. A. Zakharova, и E. V. Vekhter, «Methods of solving problems of data analysis using analytical visual models», Sci. Vis., V. 9, Is. 4, P. 78–88, 2017, doi: 10.26583/sv.9.4.08.
2. Chen, «Science Mapping: A Systematic Review of the Literature», J. Data Inf. Sci., V. 2, Is. 2, P. 1–40, 2017, doi: 10.1515/jdis-2017-0006.
3. V. Shklyar, A. A. Zakharova, E. V. Vekhter, и A. J. Pak, «Visual modeling in an analysis of multidimensional data», J. Phys. Conf. Ser., V. 5, Is. 1, P. 125–128, 2018, doi: 10.1088/1742-6596/944/1/012127.
4. J. F. Rodrigues Jr., L. A. S. Romani, A. J. M. Traina, и C. Traina Jr., «Combining Visual Analytics and Content Based Data Retrieval Technology for Efficient Data Analysis», 4th International Conference Information Visualisation, IEEE, 2010, P. 61–67. doi: 10.1109/IV.2010.101.
5. A.E. Bondarev, V.A. Galaktionov. Generalized Computational Experiment and Visual Analysis of Multidimensional Data (2019). Scientific Visualization 11.4: 102 - 114, DOI: 10.26583/sv.11.4.09.

6. Vieira, P. Parsons, и V. Byrd, «Visual learning analytics of educational data: A systematic literature review and research agenda», *Comput. Educ.*, 2018, doi: 10.1016/j.compedu.2018.03.018.
7. H. Chen и др., «Uncertainty-Aware Multidimensional Ensemble Data Visualization and Exploration», *IEEE Trans. Vis. Comput. Graph.*, V. 21, Is. 9, P. 1072–1086, 2015, doi: 10.1109/TVCG.2015.2410278.
8. V. L. Averbukh и др., «Virtual reality as an instrument of computer visualization», в 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), 2019, P. 0786–0792. doi: 10.1109/SIBIRCON48586.2019.8957854.
9. M. A. Yalçın, N. Elmqvist, и B. B. Bederson, «Cognitive Stages in Visual Data Exploration», в *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16*, 2016. doi: 10.1145/2993901.2993902.
10. V. N. Kasyanov, *Methods and tools for information visualization on the basis of attributed hierarchical graphs with ports*, *Siberian Aerospace Journal*. 2023, Vol. 24, No. 1, P. 8–17. Doi: 10.31772/2712-8970-2023-24-1-8-17.
11. Batch и N. Elmqvist, «The Interactive Visualization Gap in Initial Exploratory Data Analysis», *IEEE Trans. Vis. Comput. Graph.*, V. 24, Is. 1, P. 278–287, 2018, doi: 10.1109/TVCG.2017.2743990.
12. H. Guo, S. R. Gomez, C. Ziemkiewicz, и D. H. Laidlaw, «A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights», *IEEE Trans. Vis. Comput. Graph.*, V. 22, Is. 1, P. 51–60, 2016, doi: 10.1109/TVCG.2015.2467613.
13. R. Pienta и др., «VIGOR: Interactive Visual Exploration of Graph Query Results», *IEEE Trans. Vis. Comput. Graph.*, т. 24, вып. 1, сс. 215–225, янв. 2018, doi: 10.1109/TVCG.2017.2744898.
14. R. J. Crouser, L. Franklin, A. Endert, и K. Cook, «Toward Theoretical Techniques for Measuring the Use of Human Effort in Visual Analytic Systems», *IEEE Trans. Vis. Comput. Graph.*, 2017, doi: 10.1109/TVCG.2016.2598460.
15. R. Akhmadeeva, Y. A. Zagorulko, и D. I. Mouromtsev, «Ontology-Based Information Extraction for Populating the Intelligent Scientific Internet Resources», в *Knowledge Engineering and Semantic Web*, A.-C. Ngonga Ngomo и P. Křemen, Ред., Cham: Springer International Publishing, 2016, P. 119–128.
16. V. Shklyar, A. A. Zakharova, E. V. Vekhter, и D. A. Zavyalov, «Visual detection of internal patterns in the empirical data», в *Communications in Computer and Information Science*, Volgograd, Russia: Springer Verlag, 2017, P. 215–230. doi: 10.1007/978-3-319-65551-2_16.
17. Chen и M. Song, «Visualizing a field of research: A methodology of systematic scientometric reviews», *PloS One*, V. 14, Is. 10, с. e0223994, 2019.
18. Ware, *Foundation for a Science of Data Visualization*. 2012. doi: 10.1016/B978-0-12-381464-7.00001-6.
19. Meirelles, «Diagramming as A Strategy For Solving Graphic Design Problems», Просмотрено: 29 февраль 2020 г. [Онлайн]. Доступно на: https://www.academia.edu/2749924/Diagramming_As_A_Strategy_For_Solving_Graphic_Design_Problems
20. S. Nison, *Japanese Candlestick Charting Techniques*. 2024.
21. J. H. Larkin и H. A. Simon, «Why a Diagram is (Sometimes) Worth Ten Thousand Words», *Cogn. Sci.*, V. 11, Is. 1, P. 65–100, 1987, doi: 10.1111/j.1551-6708.1987.tb00863.x.
22. Branchini, U. Savardi, и I. Bianchi, «Productive thinking: The role of perception and perceiving opposition», *Gestalt Theory*, V. 37, Is. 1, P. 7–24, 2015
23. Branchini, U. Savardi, и I. Bianchi, «Productive thinking: The role of perception and perceiving opposition», *Gestalt Theory*, V. 37, Is. 1, P. 7–24, 2015

24. J. Choi, S. Jung, D. G. Park, J. Choo, и N. Elmqvist, «Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization», Comput. Graph. Forum, V. 38, Is. 3, P. 249–260, 2019, doi: 10.1111/cgf.13686.