

СПОСОБЫ ВИЗУАЛИЗАЦИИ МНОГОМЕРНЫХ ДАННЫХ

Бородачева Т. Д.¹, Шкляр А.В.²

¹ТПУ ИШИТР ОАР, гр. 8Д11, e-mail: tdb3@tpu.ru

²ТПУ ИШИТР ОАР, доцент, e-mail: shklyarav@tpu.ru

Аннотация

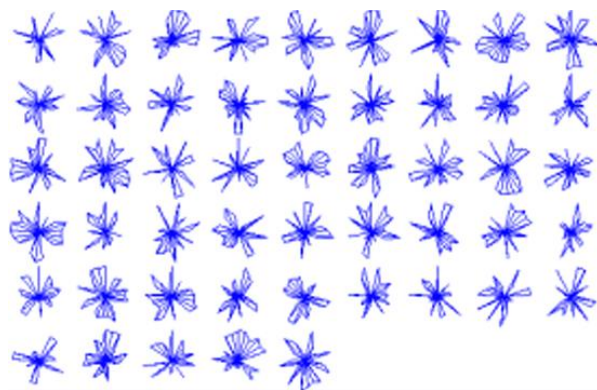
В работе рассматривается проблема визуализации сложных для анализа человека многомерных данных, а также предлагается альтернативный подход к демонстрации характеристик систем и объектов через использование единого визуального образа как многомерного массива.

Ключевые слова: многомерные данные, массив данных, визуализация, сопоставление, анализ, информация.

Введение

В современном мире наблюдается стремительный рост объема данных, что делает актуальным вопрос эффективного анализа и интерпретации многомерной информации. Визуализация такой информации упрощает её исследование за счет демонстрации множества переменных и взаимосвязей в виде упрощенных зрительных образов.

Существующие способы визуализации многомерных данных служат инструментом, который помогает преобразовать множество данных в более наглядные графические представления, доступные для восприятия и анализа, однако, каждый из них имеет ряд ограничений, связанных с особенностями сферы их применения [1]. Способы, фиксирующие в едином визуальном представлении большое количество данных об объекте, зачастую трудно применимы для задач сопоставления и выявления скрытых закономерностей. Помимо этого, получаемые графические образы являются специфичными, что определяет необходимость наличия узконаправленных знаний для их интерпретации и оценки. Подобными способами визуализации являются: диаграммы-радары, диаграммы с параллельными координатами [2]. На рисунке 1 представлены примеры таких диаграмм.



а)



б)

Рис. 1. Пример визуализации данных с помощью:

а – множества диаграмм-радаров, б – диаграммы с параллельными координатами

Таким образом, существует потребность в создании более универсальной системы, использующей для демонстрации данных простые элементы, образующие цельный, легкий для восприятия, образ.

Основная часть

Предлагаемое решение основано на методике «Лица Чернова», где элементы, фиксирующие разные характеристики объекта, складываются в единый визуальный образ. Основная идея способа заключается в использовании различных черт лица как элементов, содержащих информацию о значении переменных. Таким образом, каждое лицо становится уникальным представителем конкретного объекта в многомерном пространстве [3]. Преимущество метода заключается в его свойстве быстро передавать информацию за счет интуитивного восприятия лиц людьми, что облегчает интерпретацию данных. На рисунке 2 представлен пример использования «Лиц Чернова» для демонстрации уровня социально-значимых параметров, таких как доля населения за чертой бедности, распространённость высшего образования, процент безработицы и кол-во взрослых с ожирением по Соединенным Штатам Америки.



Рис. 2. Пример визуализации данных «Лицами Чернова»

Несмотря на наглядность и удобство данного способа визуализации, у него есть и особенности, не всегда положительно влияющие на процесс анализа информации. Так, число объектов доступных для оценки и сопоставления очень ограничено. При увеличении количества сравниваемых объектов, появляется необходимость в масштабировании общей схемы. В такой ситуации черты лица становятся менее распознаваемыми, что не способствует удобному сопоставлению отдельных показателей разных объектов.

Лицо является интуитивно понятным для человека образом, что эффективно применяется в описываемом способе визуализации. При этом, лица используются для демонстрации данных, напрямую связанных с человеком. Так, можно предположить, что образ лица не является универсальным, и плохо подходит для работы с данными, не описывающими состояние человека.

Колористическое решение представляет собой один цвет, тон которого изменяется от светлого к более тёмному. Цвет каждого лица показывает общее состояние объекта, резюмируя все его характеристики. Чем темнее цвет лица, тем хуже ситуация с состоянием населения. Данный прием хорошо подходит для задачи выявления объектов с лидирующей позицией, однако, не удобен при необходимости группировки объектов.

Разработанный способ визуализации многомерных данных учитывает описанные выше особенности способа визуализации «Лицами Чернова». За основу решения взят простой для восприятия образ цветка, состоящий из множества элементов, вариации которых содержат определенную информацию. Цветок может состоять из следующих элементов: лепестков – это наиболее заметная часть цветка; тычинок; стеблей; листьев. Каждый из элементов может быть представлен в разном количестве, размере и цвете. Определённое сочетание элементов и их цветов может привести к изменению восприятия образа, сделать его более привлекательным

или отталкивающим. Цветовое решение и форма элементов также могут быть адаптированы под задачи группировки и выбора среди множества.

Образ цветка уместен для подачи информации различной тематики, так как может быть интерпретирован в качестве следующих метафор:

1. Роста и развития.
2. Сезонности и перемен.
3. Надежды и возрождение.
4. Единства и разнообразия.

Для демонстрации применения разработанного способа далее в работе используются данные информационно-аналитической системы городского развития – индекс качества жизни в городах России. [4]. Индекс качества оценивает города России по стобалльной системе в одиннадцати категориях, среди которых присутствуют: доход и работа, здоровье, благоустройство, природно-экологические условия, безопасность. Каждой категории в системе визуализации присвоен свой элемент или его свойство:

1. Уровень дохода и занятости отражает количество лепестков цветка.
2. Уровень здоровья отражает размер тычинки.
3. Уровень благоустройства отражает тональность цвета цветка.
4. Природно-экологические условия отражает цвет цветка.
5. Уровень безопасности отражает форма цветка.

Для удобства демонстрации данных, стобалльная система оценивания каждого из критериев преобразована в пятиступенчатую шкалу. Таким образом, каждый из визуальных элементов иконки имеет 5 видов представления, изменяющихся по одному из свойств (размеру, форме, цвету).

На рисунке 3 продемонстрирована цветовая палитра, состоящая из градации пяти цветов (от фиолетового к зеленому) и их пяти разных тонов. Изменение цвета означает изменение уровня природно-экологических условий. Зеленый цвет – означает самый высокий уровень природно-экологических условий (от 81 до 100 %), а фиолетовый цвет – самый низкий (от 0 до 20 %). При этом, самый тусклый оттенок каждого цвета обозначает самый низкий уровень благоустройства, а самый яркий – самый высокий.



Рис. 3. Визуализация изменения уровня природно-экологических условий и благоустройства

На рисунке 4 приведено изменение формы цветка. Каждая последующая фигура, начиная с первой, имеет большее количество лепестков, чем предыдущая. Данная закономерность отражает изменение значения уровня дохода и занятости населения.

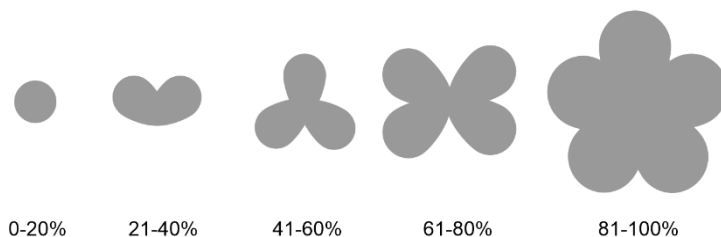


Рис. 4. Визуализация изменения уровня работы и занятости

На рисунке 5 показано изменение размера тычинки цветка. Чем больше размер тычинки, тем выше уровень здравоохранения в описываемом городе.



Рис. 5. Визуализация изменения уровня здравоохранения

На рисунке 6 изображен пример изменения формы лепестков цветка, отражающий уровень безопасности в городе. Чем острее и тоньше лепестки – тем более опасным является город. Округлая форма лепестков свидетельствует о высоком уровне безопасности.



Рис. 6. Пример визуализации изменения уровня безопасности

На основе статистических данных о качестве жизни в городе Томске был составлен визуальный образ, элементы которого отображают соответствующие числовые значения по принципу, описанному выше. Изображение модели качества жизни в городе Томске показано на рисунке 7, б. На рисунке 7, а отображена та же информация, но в виде таблицы. Показатели по городу отмечены галочкой в соответствующей каждому параметру строке.

	0-20	21-40	41-60	61-80	81-100
Работа			✓		
Здоровье			✓		
Благоустройство				✓	
Экология			✓		
Безопасность				✓	

а)

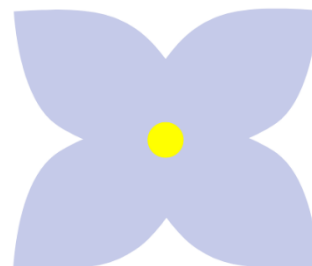


б)

Рис. 7. Пример визуализации данных о качестве жизни в городе Томске с помощью: а – таблицы; б – авторского способа визуализации многомерных данных

За счет разного подхода к визуализации многомерных данных, графические изображения на рисунках 7, а и 7, б, считаются с разной скоростью. Изображение 7, б представляет собой единый образ, и по этой причине воспринимается и анализируется быстрее. Однако, его использование удобнее в системе из подобных образов, содержащих разные значения переменных. В таких условиях становится возможным сопоставление разных объектов, а также облегчается анализ свойств каждого из них по отдельности благодаря считываемой разнице между ними. На рисунке 8, а представлена визуализация данных о качестве жизни в городе Братске в виде таблицы, а на рисунке 8, б в виде цветка.

	0-20	21-40	41-60	61-80	81-100
Работа				✓	
Здоровье		✓			
Благоустройство		✓			
Экология		✓			
Безопасность			✓		



*Рис. 8. Пример визуализации данных о качестве жизни в городе Братске с помощью:
а – таблицы; б – авторского способа визуализации многомерных данных*

При сравнении визуализаций на рисунке 7, б и 8, б, можно быстро заметить существенные различия по форме, размеру и цвету элементов, что говорит и о разнице показателей качества жизни в этих городах.

Заключение

Разработанный способ является альтернативным подходом к визуализации многомерных данных. За счет универсальности заложенного визуального образа, он может быть применен для представления данных разной тематики, что делает его универсальным и удобным в использовании. Вариативность элементов позволяет адаптировать способ под задачи как группировки, так и поиска выделяющегося объекта. С учетом роста количества информации, способ является актуальным и может быть применен для эффективного анализа и обработки данных.

Список использованной литературы:

1. Васильев В.Р., Волобой А.Г., Вьюкова Н.И., Галактионов В.А. Контекстная визуализация пространственных данных // Препринты ИПМ им. М.В. Келдыша. – 2004. – № 56. – 23 с. – [Электронный ресурс]. – URL: keldysh.ru/pages/cgraph/articles/dep20/vis_pp.pdf (дата обращения: 20.03.2025).
2. Зиновьев А.Ю. Визуализация многомерных данных. Красноярск : Изд-во КГТУ. – 2000. – 168 с.
3. Осадчая И.А., Берестнева О.Г., Немеров Е.В. Анализ многомерных медицинских данных с помощью пиктографиков «Лица Чернова». – [Электронный ресурс] – URL: cyberleninka.ru/article/n/analiz-mnogomernyh-meditsinskih-dannyh-s-pomoschyu-piktografikov-litsa-chernova (дата обращения 20.03.2025).
4. Индекс качества жизни. – [Электронный курс] – URL: citylifeindex.ru/ (дата обращения 20.03.2025).